

Domestic mite Bt11	12	116	8.7	1017	22	AAB02246
Putative P. abyssi protein with Rho P	13	116	8.7	1177	22	AAB96721
Male enhanced anti-Mouse male enhance	14	115.6	8.7	1372	19	AAW56473
Novel human diagno	15	115	8.6	1325	18	AAW19510
Drosophila melanog	16	115	8.6	1325	20	AAB94391
Human Protein seq	17	113.5	8.5	881	22	ABG05240
Corn MFPL1. Zea ma	18	113.5	8.5	881	22	ABG20238
Novel human diagno	19	112.5	8.4	622	22	ABG26216
Drosophila melanog	20	112	8.4	789	22	AAB95410
Novel human diagno	21	111.5	8.4	672	21	AAB21233
Enterococcus faeca	22	111.5	8.4	1192	22	AAU35210
Protein with Rho P	23	111.5	8.4	1388	19	AAW56475
Membrane antigen P	24	111	8.3	334	15	AAY51275
Human glial fibril	25	111	8.3	433	19	AAV2975
Partial protein en	26	109.5	8.2	398	21	AAB21228
Novel human diagno	27	109	8.2	2415	22	ABG20779
Novel human diagno	28	108.5	8.1	513	22	ABG08133
Drosophila melanog	29	108.5	8.1	970	22	ABD70846
Drosophila melanog	30	108.5	8.1	1081	21	AAG50490
Arabidopsis thalia	31	108.5	8.1	1197	21	AAG50489
Drosophila melanog	32	108.5	8.1	1690	22	ABB61144
Drosophila melanog	33	108.5	8.1	1690	22	ABB61173
Drosophila melanog	34	108.5	8.1	2013	22	ABB63322
Breast and ovarian	35	108	8.1	430	21	AAC25548
Human serollectin.	36	108	8.1	458	21	AAG25547
Arabidopsis thalia	37	108	8.1	473	21	AAG25546
Novel human diagno	38	108	8.1	776	22	ABG05279
Novel human diagno	39	108	8.1	777	22	ABG21257
Human serollectin.	40	107.5	8.1	433	21	AAB57557
Amino acid sequenc	41	107.5	8.1	469	18	AAW23820
Human keratin KERT	42	107.5	8.1	469	21	AAW61289
Receptor for hyalu	43	107.5	8.1	546	21	AAV52397
RHAMM 1-2 isoform	44	107.5	8.1	606	17	AAR93673
Chlamydia trachomatis infection-specific protein InCa.	45	107.5	8.1	631	17	AAR93675
Chlamydia trachomatis.	46	108	8.1			
W0953948-A1.	47	108	8.1			
PD 28-OCT-1999.	48	108	8.1			
XX 20-APR-1999;	49	108	8.1			
XX 20-APR-1998;	50	108	8.1			
XX 21-APR-1998;	51	108	8.1			
PR 21-MAY-1998;	52	108	8.1			
PR 22-MAY-1998;	53	108	8.1			
XX 01-FEB-2000 (first entry)	54	108	8.1			
DE Chlamydia trachomatis infection-specific protein InCa.	55	108	8.1			
XX InCa; infection; vaccine; therapy; diagnosis.	56	108	8.1			
OS Chlamydia trachomatis.	57	108	8.1			
XX W0953948.	58	108	8.1			
XX PN W0953948-A1.	59	108	8.1			
XX PD 28-OCT-1999.	60	108	8.1			
XX XX 20-APR-1999;	61	108	8.1			
XX XX 20-APR-1998;	62	108	8.1			
XX PR 21-APR-1998;	63	108	8.1			
XX PR 22-MAY-1998;	64	108	8.1			
XX PR 22-MAY-1998;	65	108	8.1			
XX PR 22-MAY-1998;	66	108	8.1			
XX PR 22-MAY-1998;	67	108	8.1			
XX PR 22-MAY-1998;	68	108	8.1			
XX PR 22-MAY-1998;	69	108	8.1			
XX PR 22-MAY-1998;	70	108	8.1			
XX PR 22-MAY-1998;	71	108	8.1			
XX PR 22-MAY-1998;	72	108	8.1			
XX PR 22-MAY-1998;	73	108	8.1			
XX PR 22-MAY-1998;	74	108	8.1			
XX PR 22-MAY-1998;	75	108	8.1			
XX PR 22-MAY-1998;	76	108	8.1			
XX PR 22-MAY-1998;	77	108	8.1			
XX PR 22-MAY-1998;	78	108	8.1			
XX PR 22-MAY-1998;	79	108	8.1			
XX PR 22-MAY-1998;	80	108	8.1			
XX PR 22-MAY-1998;	81	108	8.1			
XX PR 22-MAY-1998;	82	108	8.1			
XX PR 22-MAY-1998;	83	108	8.1			
XX PR 22-MAY-1998;	84	108	8.1			
XX PR 22-MAY-1998;	85	108	8.1			
XX PR 22-MAY-1998;	86	108	8.1			
XX PR 22-MAY-1998;	87	108	8.1			
XX PR 22-MAY-1998;	88	108	8.1			
XX PR 22-MAY-1998;	89	108	8.1			
XX PR 22-MAY-1998;	90	108	8.1			
XX PR 22-MAY-1998;	91	108	8.1			
XX PR 22-MAY-1998;	92	108	8.1			
XX PR 22-MAY-1998;	93	108	8.1			
XX PR 22-MAY-1998;	94	108	8.1			
XX PR 22-MAY-1998;	95	108	8.1			
XX PR 22-MAY-1998;	96	108	8.1			
XX PR 22-MAY-1998;	97	108	8.1			
XX PR 22-MAY-1998;	98	108	8.1			
XX PR 22-MAY-1998;	99	108	8.1			
XX PR 22-MAY-1998;	100	108	8.1			
XX PR 22-MAY-1998;	101	108	8.1			
XX PR 22-MAY-1998;	102	108	8.1			
XX PR 22-MAY-1998;	103	108	8.1			
XX PR 22-MAY-1998;	104	108	8.1			
XX PR 22-MAY-1998;	105	108	8.1			
XX PR 22-MAY-1998;	106	108	8.1			
XX PR 22-MAY-1998;	107	108	8.1			
XX PR 22-MAY-1998;	108	108	8.1			
XX PR 22-MAY-1998;	109	108	8.1			
XX PR 22-MAY-1998;	110	108	8.1			
XX PR 22-MAY-1998;	111	108	8.1			
XX PR 22-MAY-1998;	112	108	8.1			
XX PR 22-MAY-1998;	113	108	8.1			
XX PR 22-MAY-1998;	114	108	8.1			
XX PR 22-MAY-1998;	115	108	8.1			
XX PR 22-MAY-1998;	116	108	8.1			
XX PR 22-MAY-1998;	117	108	8.1			
XX PR 22-MAY-1998;	118	108	8.1			
XX PR 22-MAY-1998;	119	108	8.1			
XX PR 22-MAY-1998;	120	108	8.1			
XX PR 22-MAY-1998;	121	108	8.1			
XX PR 22-MAY-1998;	122	108	8.1			
XX PR 22-MAY-1998;	123	108	8.1			
XX PR 22-MAY-1998;	124	108	8.1			
XX PR 22-MAY-1998;	125	108	8.1			
XX PR 22-MAY-1998;	126	108	8.1			
XX PR 22-MAY-1998;	127	108	8.1			
XX PR 22-MAY-1998;	128	108	8.1			
XX PR 22-MAY-1998;	129	108	8.1			
XX PR 22-MAY-1998;	130	108	8.1			
XX PR 22-MAY-1998;	131	108	8.1			
XX PR 22-MAY-1998;	132	108	8.1			
XX PR 22-MAY-1998;	133	108	8.1			
XX PR 22-MAY-1998;	134	108	8.1			
XX PR 22-MAY-1998;	135	108	8.1			
XX PR 22-MAY-1998;	136	108	8.1			
XX PR 22-MAY-1998;	137	108	8.1			
XX PR 22-MAY-1998;	138	108	8.1			
XX PR 22-MAY-1998;	139	108	8.1			
XX PR 22-MAY-1998;	140	108	8.1			
XX PR 22-MAY-1998;	141	108	8.1			
XX PR 22-MAY-1998;	142	108	8.1			
XX PR 22-MAY-1998;	143	108	8.1			
XX PR 22-MAY-1998;	144	108	8.1			
XX PR 22-MAY-1998;	145	108	8.1			
XX PR 22-MAY-1998;	146	108	8.1			
XX PR 22-MAY-1998;	147	108	8.1			
XX PR 22-MAY-1998;	148	108	8.1			
XX PR 22-MAY-1998;	149	108	8.1			
XX PR 22-MAY-1998;	150	108	8.1			
XX PR 22-MAY-1998;	151	108	8.1			
XX PR 22-MAY-1998;	152	108	8.1			
XX PR 22-MAY-1998;	153	108	8.1			
XX PR 22-MAY-1998;	154	108	8.1			
XX PR 22-MAY-1998;	155	108	8.1			
XX PR 22-MAY-1998;	156	108	8.1			
XX PR 22-MAY-1998;	157	108	8.1			
XX PR 22-MAY-1998;	158	108	8.1			
XX PR 22-MAY-1998;	159	108	8.1			
XX PR 22-MAY-1998;	160	108	8.1			
XX PR 22-MAY-1998;	161	108	8.1			
XX PR 22-MAY-1998;	162	108	8.1			
XX PR 22-MAY-1998;	163	108	8.1			
XX PR 22-MAY-1998;	164	108	8.1			
XX PR 22-MAY-1998;	165	108	8.1			
XX PR 22-MAY-1998;	166	108	8.1			
XX PR 22-MAY-1998;	167	108	8.1			
XX PR 22-MAY-1998;	168	108	8.1			
XX PR 22-MAY-1998;	169	108	8.1			
XX PR 22-MAY-1998;	170	108	8.1			
XX PR 22-MAY-1998;	171	108	8.1			
XX PR 22-MAY-1998;	172	108	8.1			
XX PR 22-MAY-1998;	173	108	8.1			
XX PR 22-MAY-1998;	174	108	8.1			
XX PR 22-MAY-1998;	175	108	8.1			
XX PR 22-MAY-1998;	176	108	8.1			
XX PR 22-MAY-1998;	177	108	8.1			
XX PR 22-MAY-1998;	178	108	8.1			
XX PR 22-MAY-1998;	179	108	8.1			
XX PR 22-MAY-1998;	180	108	8.1			
XX PR 22-MAY-1998;	181	108	8.1			
XX PR 22-MAY-1998;	182	108	8.1			
XX PR 22-MAY-1998;	183	108	8.1			
XX PR 22-MAY-1998;	184	108	8.1			
XX PR 22-MAY-1998;	185	108	8.1			
XX PR 22-MAY-1998;	186	108	8.1			
XX PR 22-MAY-1998;	187	108	8.1			
XX PR 22-MAY-1998;	188	108	8.1			
XX PR 22-MAY-1998;	189	108	8.1			
XX PR 22-MAY-1998;	190	108	8.1			
XX PR 22-MAY-1998;	191	108	8.1			
XX PR 22-MAY-1998;	192	108	8.1			
XX PR 22-MAY-1998;	193	108	8.1			
XX PR 22-MAY-1998;	194	108	8.1			
XX PR 22-MAY-1998;	195	108	8.1			
XX PR 22-MAY-1998;	196	108	8.1			
XX PR 22-MAY-1998;	197	108	8.1			
XX PR 22-MAY-1998;	198	108	8.1			
XX PR 22-MAY-1998;	199	108	8.1			
XX PR 22-MAY-1998;	200	108	8.1			
XX PR 22-MAY-1998;	201	108	8.1			
XX PR 22-MAY-1998;	202	108	8.1			
XX PR 22-MAY-1998;	203	108	8.1			
XX PR 22-MAY-1998;	204	108	8.1			
XX PR 22-MAY-1998;	205	108	8.1			
XX PR 22-MAY-1998;	206	108	8.1			
XX PR 22-MAY-1998;	207	108	8.1			
XX PR 22-MAY-1998;	208	108	8.1			
XX PR 22-MAY-1998;	209	108	8.1			
XX PR 22-MAY-1998;	210	108	8.1			
XX PR 22-MAY-1998;	211	108	8.1			
XX PR 22-MAY-1998;	212	108	8.1			
XX PR 22-MAY-1998;	213	108	8.1			
XX PR 22-MAY-1998;	214	108	8.1			
XX PR 22-MAY-1998;	215	108	8.1			
XX PR 22-MAY-1998;	216	108	8.1			
XX PR 22-MAY-1998;	217	108	8.1			
XX PR 22-MAY-1998;	218	108	8.1			
XX PR 22-MAY-1998;	219	108	8.1			
XX PR 22-MAY-1998;						

PS Claim 5; Page 45-46; 56pp; English.

XX This sequence represents a novel infection-specific protein Inca of Chlamydia trachomatis LGV-134 (serotype L2). Inca is found in the inclusion membrane of infected cells. It is associated primarily with the vegetative elementary body form. The invention includes a vaccine directed against the reticulate body form of Chlamydia comprising 1 or more infection-specific proteins (see AAY32170-78), including Inca, IncB and IncC; methods of using and producing such a vaccine; methods for detection of infection-specific antibodies or antigens in a biological specimen; and method of using therapeutic agents specifically directed against infection-specific peptides, or the genes that code for such peptides, to treat chlamydial infection.

CC XX Sequence 273 AA;

CC SQ Score 100.0%; Score 1333; DB 20; Length 273;

CC Best Local Similarity 100.0%; Pred. No. 2, 6e-106; Indels 0; Gaps 0;

CC Matches 273; Conservative 0; Mismatches 0; Delins 0; Gaps 0;

QY 1 MTPPTLIVPPSPPPSPSANSRVPQPSLMDKIKKIAAATASLILIGTIGFLALLGHLYGFL 60

CC 1 mtpptlivppspapsysanrvqps.lmdkkiaaaatasliligtigflallghlygfl 60

DB 61 IAPQITIVLLALFTTSLAGNALLQKTAANLHYQDLOREVGSIKEINPMLSVLQKEFHL 120

CC 61 iapqitivllalfttslagnallyqktanlahlyqdloregvsiikeinmlsvlqkefhl 120

QY 121 SKEFATTSDLSAVSQDFYSCLQGRFDNYKFGESLDEYNSTEEMRKLESQETIADLK 180

CC 121 skefattsdlsavsqdfysclqgrfdnykfgesldeynsteemrklesqetiadlk 180

DB 121 skefattsdlsavsqdfysclqgrfdnykfgesldeynsteemrklesqetiadlk 180

QY 181 SVASLREEIRFLTPLAAEVRLAHNOESTATAEELKTIRDSLRDEIGOISQLSKTLTSQ 240

CC 181 svaslreeirfltplaaevrlahngestataelktirdsirdeigoisqlsktltsq 240

DB 181 svaslreeirfltplaaevrlahngestataelktirdsirdeigoisqlsktltsq 240

QY 241 IALORKESSDCSQTRELTSPREASAPTKSS 273

CC 241 ialdrkessdcsgretlssprksapskss 273

DB 241 ialdrkessdcsgretlssprksapskss 273

---

RESULT 2 AAY37087 standard; Protein: 195 AA.

XX DT 07-OCT-1999 (first entry)

DE Amino acid sequence of a Chlamydia trachomatis protein.

XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma; paratrachoma; inclusion conjunctivitis; genital disease; perinephritis; iongonococcal urethritis; epidymitis; cervicitis; salpingitis; bartholinitis; pneumopathy; venereal lymphogranulomatosis.

KW XX Chlamydia trachomatis.

OS XX WO9928475-A2.

PN PD 10-JUN-1999.

XX PF 27-NOV-1998; 98WO-IB01939.

XX PR 04-NOV-1998; 98US-0107077.

PR 28-NOV-1997; 97FR-0015041.

PR 17-DEC-1997; 97FR-0016334.

XX PA (GEST ) GENSET.

XX

PI Griffais R;

XX WPI: 1999-371125/31.

DR Genome sequence of Chlamydia trachomatis.

XX Disclosure; Page 892-893; 1755pp; English.

PS AAY36754-Y7949 are encoded by open reading frames (ORFs) of the genome of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences can also be used to control growth of the microorganism. Chlamydia trachomatis is responsible for a large number of diseases, e.g. eye diseases such as conventional trachoma, nonendemic trachoma, paratrachoma, and inclusion conjunctivitis; genital diseases such as nongonococcal urethritis, epidymitis, cervicitis; pneumopathy in infants; perinephritis, bartholinitis; pneumopathy in breast feeding infants; and venereal lymphogranulomatosis. The polypeptides of the invention may be of use in treating these diseases.

XX Sequence 195 AA;

QY Query Match 63.0%; Score 840; DB 20; Length 195;

CC Best Local Similarity 91.5%; Pred. No. 3, 2e-64;

CC Matches 173; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 64 QTTIVLAFLFTSLAGNALLQKTAANLHYQDLOREVGSIKEINPMLSVLQKEFHL 123

CC 64 kftenvlalfttslagnallyqktanlahlyqdloregvsiikeinmlsvlqkefhl 63

DB 124 FATTSDLSAVSQDFYSCLQGRFDNYKFGESLDEYNSTEEMRKLESQETIADLK 183

CC 124 fattskdlssavsqdfysclqgrfdnykfgesldeynsteemrklesqetiadlk 123

DB 64 fattskdlssavsqdfysclqgrfdnykfgesldeynsteemrklesqetiadlk 183

QY 184 STREERIRFLTPLAAEVRLAHNOESTATAEELKTIRDSLRDEIGOISQLSKTLTSQ 243

CC 184 streeerirfltplaaevrlahngestataelktirdsirdeigoisqlsktltsq 243

DB 124 streeerirfltplaaevrlahngestataelktirdsirdeigoisqlsktltsq 183

QY 244 QRKESSDC 252

CC 244 qrkessc 252

DB 184 ynekrqac 192

---

RESULT 3 AAY37086

ID AAY37086 standard; Protein: 81 AA.

XX AC AAY37086;

XX DT 07-OCT-1999 (first entry)

DE Amino acid sequence of a Chlamydia trachomatis protein.

XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma; paratrachoma; inclusion conjunctivitis; genital disease; perinephritis; iongonococcal urethritis; epidymitis; cervicitis; salpingitis; bartholinitis; pneumopathy; venereal lymphogranulomatosis.

KW XX Chlamydia trachomatis.

OS XX WO9928475-A2.

PN PD 10-JUN-1999.

XX PF 27-NOV-1998; 98WO-IB01939.

XX PR 04-NOV-1998; 98US-0107077.

PR 28-NOV-1997; 97FR-0015041.

PR 17-DEC-1997; 97FR-0016334.

XX PA (GEST ) GENSET.

XX

PI Griffais R;  
 XX WPI: 1999-371125/31.  
 PT Genome sequence of Chlamydia trachomatis  
 XX Disclosure; Page 892; 175pp; English.  
 PS AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome of Chlamydia trachomatis (see AA01423). The polypeptides can be used as vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences can also be used to control growth of the microorganism. Chlamydia trachomatis is responsible for a large number of diseases, e.g. eye diseases such as conventional trachoma, nonendemic trachoma, paratrachoma, and inclusion conjunctivitis; genital diseases such as perigenital urethritis, epididymitis, cervicitis, salpingitis, per-hepatitis, bartholinitis; pneumopathy in breast feeding infants; and venereal lymphogranulomatosis. The polypeptides of the invention may be of use in treating these diseases.  
 XX Sequence 81 AA;

Query Match Score 23.8%; Best Local Similarity 98.5%; Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MTPPTLIVPPSPPABPSYANRPQPSLMDIKKIAAASLLIGTIGFLALIGHVGL 60  
 Db 16 mttptlivtpbssppabpsysanrvpqpslmdikkiaaaslligtlflalighvgl 75  
 Qy 61 IAPQI 65  
 Db 76 iapqi 80

Query Match Score 23.17; Best Local Similarity 97.2e-20; Matches 61; Conservative 0; Mismatches 1; Indels 1; Gaps 0;  
 Qy 1 MTPPTLIVPPSPPABPSYANRPQPSLMDIKKIAAASLLIGTIGFLALIGHVGL 60  
 Db 16 mttptlivtpbssppabpsysanrvpqpslmdikkiaaaslligtlflalighvgl 75  
 Qy 61 IAPQI 65  
 Db 76 iapqi 80

RESULT 4  
 AAY32173 ID AAY32173 standard; protein; 355 AA.  
 XX DT 01-FEB-2000 (first entry)  
 XX DE Chlamydia psittaci infection-specific protein IncA.  
 XX KW IncA; infection; vaccine; therapy; diagnosis.  
 OS Chlamydia psittaci.  
 W0953948-A1.  
 PD 28-OCT-1999.  
 XX PF 20-APR-1999; 99WO-US08744.  
 XX PR 20-APR-1998; 98US-0082438.  
 PR 21-APR-1998; 98US-0082588.  
 PR 22-MAY-1998; 98US-0086450.  
 PA (UVR-) UNIV OREGON STATE.  
 PI Rockey DD, Bannantine JP;  
 DR WPI: 1999-633904/54.  
 DR N-PSDB: AA234587.  
 PT Novel bacterial infection specific proteins for treating and diagnosing chlamydial infections -  
 XX PS Claim 5; Page 39-40; 56pp; English.  
 CC This sequence represents novel infection-specific protein IncA of

CC Chlamydia psittaci strain GPIC. IncA is found in the inclusion membrane of infected cells. It is associated primarily with the vegetative reticulate body form of Chlamydia rather than with the refractive elementary body form. The invention includes: a vaccine directed against the reticulate body form of Chlamydia comprising 1 or more infection-specific proteins (see AAY32170-78), including IncA, IncB and IncC; methods of using and producing such a vaccine; methods for detection of infection-specific antibodies or antigens in a biological specimen; and a method of using therapeutic agents specifically directed against infection-specific peptides, or the genes that code for such peptides, to treat chlamydial infection.

CC Sequence 355 AA;

Query Match Score 12.28; Best Local Similarity 21.4%; Matches 63; Conservative 61; Mismatches 111; Indels 59; Gaps 8;  
 Qy 8 VIPPSPPPAPSYSANRY- PQQPSLMDKIKKIAAASLLIGTIGFLALIGHVGC 58  
 Db 32 lipisteaepssaaavgaktaeppgrspqrlqcykvkaalaafvvgiaalvclyg 91  
 Qy 59 FLIAPOQTIVLLAFTITSLA-----GNALYLOKTAANLHYQLDQR-----EV 100  
 Db 92 svistpsl1mlamlnvsvivitardgtspqvrrr--hmikcqgqqfgeentrlhtav 147  
 Qy 101 GSLKRINFMLS--YIQLKEFLH----LSKEFATTSKDLASAVSQDFYSCLOGFRDNYKGPE 153  
 Db 148 enlkavnvelseqinqkqtrisdfgrleantdfallqslsleefsvgtkve 207  
 Qy 154 SLIDEEYKNSTEEMRKLFQSBEIIADLGKGSYASLRREIRFLPLAEVRLAHNQESTLAAI 213  
 Db 208 tm1spfeklaqs1ketsjeavqammsavtelr-----mlnalkei 250

RESULT 5  
 AAY34783 ID AAY34783 standard; protein; 397 AA.  
 XX AC AAY34783;  
 XX DT 13-SEP-1999 (first entry)  
 XX DE Chlamydia pneumoniae transmembrane protein sequence.  
 KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis; sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
 OS Chlamydia pneumoniae.  
 PN W095927105-A2.  
 XX PD 03-JUN-1999.  
 XX PF 20-NOV-1998; 98WO-1B01890.  
 XX PR 04-NOV-1998; 98US-0107078.  
 PR 21-NOV-1997; 97FR-0014673.  
 PA GEST ) GENSET.  
 XX PI Griffais R;  
 XX DR WPI; 1999-357842/30.  
 XX PT Genome sequence of Chlamydia pneumoniae

PS Page 760-761; Disclosure; 1912pp; English.  
 XX  
 CC AAV34584-Y35879 represent the proteins encoded by all the open reading frames in the complete genome (see AAX91930) of Chlamydia pneumoniae.  
 CC C. pneumoniae causes respiratory disease such as pneumonia and bronchitis and is thought to be a contributing factor in heart disease, sarcoidosis, sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The polypeptides encoded by the open reading frames of the C. pneumoniae genome (see AAV34584-Y35879) can be used in immunogenic compositions as vaccines. Vectors containing C. pneumoniae nucleotides sequences can also be used as immunogenic compositions, especially where the vector directs the expression of a neutralising epitope of C. pneumoniae.

XX Sequence 397 AA:  
 SQ 11.6%; Score 154.5; DB 20; Length 397;  
 [REDACTED] Best Local Similarity 20.8%; Pred. No. 5.2e-05; Gaps 12;  
 Matches 75; Conservative 59; Mismatches 123; Indels 103; Gaps 12;  
 Qy 3 TPTLIVIP-PSPPAPPSYSAANRVPDPSLMRKIKKIAIASLILIGTIGFLALIQLHVGF- 59  
 Db 15 tpsapnpiapptcgtptk-prsf--iekvivaklfiaatsqalgtilglsq 70  
 Qy 60 LIAPQITIVILALAFITSLAGNALYIQLQTAHLYQDLOREVGSSLKEINFLMSVLYQKEFLH 119  
 Db 71 altpgqiallviifvsvmlqigkldsiiggeirreirvsrftsendsrltv----- 123  
 Qy 120 LSKEFATTSKDLSAVSDFYSCLQGFRD--NYK----- 150  
 Qy 124 itttletekvdkalkaaqdqtlieefarnengnlkttaedleeqysklsksealealingd 183  
 Db 151 -----GFES-LLEDEYKNSTEEMRKLFQSOETIADIKGSVASLREEL 189  
 Qy 184 iqanagdaqeisselkklisgwdskvveqintsklkvlgwvqeadthvamqeqi 243  
 Db 184 RFL-----TPIAEEVRLAHNQESTAAIBELTIRD-----SLRDEIGOLS 231  
 Qy 190 RFL-----TPIAEEVRLAHNQESTAAIBELTIRD-----SLRDEIGOLS 231  
 Db 244 qalqaeilgmhngstalqksvenllyrqdqlaltrvrgellesekn1sqacalqrgeleka 303  
 Qy 232 OLSKTLTSQI-----ALQR-----KESSDLCSQIRTEFLSSPRKSASPST 270  
 Db 304 qhetlsgridamlaqneqlaeqtaekmkqeaqkraesefiacyvdrtfrtfrerpppt 363

RESULT 6  
 XX AAR72826 standard; Protein; 2482 AA.  
 AC AAR72826;  
 DT 27-FEB-1996 (first entry)  
 DE Human mitosin.  
 XX Cell cycle; M phase; mitosin; retinoblastome; mitosis; cell growth;  
 KW inhibition.  
 XX Homo sapiens.  
 CS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Region 1480..159  
 FT Region /label=internal\_repeat  
 FT Region 1660..1839  
 FT Region /label=internal\_repeat  
 PN WO9511309-A2.  
 PD 27-APR-1995.  
 XX PF 24-OCT-1994;  
 XX XX

PR 22-OCT-1993; 93US-0141239.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX Lee W. Zhu X;  
 PI DR WPI; 1995-170229/22.  
 DR N-PSDB; AAQ6851.  
 XX Purified mammalian protein mitosin and agents that bind it and inhibit its action - used to promote cell growth or to inhibit cell division and/or proliferation  
 XX  
 PS Claim 4; Fig 8B; 61pp; English.  
 XX AAR72829 is human mitosin. Mitosin is involved in the regulation of the mammalian mitotic cell cycle. Mitosin as with E2F-1 (see AAT72824) interacts with the retinoblastoma protein (the retinoblastoma tumour suppressor gene product). Mitosin is first synthesised at the G1/S boundary, it is then phosphorylated from S through M phase, and during mitosis, is closely associated with the centromeres/kinetochores at the mitotic spindle poles. Mitosin is necessary for a eukaryotic cell to enter the M phase of the mitotic cell cycle and its degradation is necessary for a cell to advance on to the next stage. Mitosin is thus useful for controlling cell growth as overexpression of mitosin prevents a cell from exiting the M phase.  
 An anti-mitosin antibody, antibody fragment or a phosphorylated mitosin mutant ( or nucleic acid encoding it) can also be used to inhibit cell division which is particularly useful for the study of the cell cycle. A further use is to control hyperproliferative cells, and so control diseases such as psoriasis and breast cancer. It can also be used to block gametogenesis of an immature gamete.  
 XX Sequence 2482 AA;  
 SQ Query Match 9.3%; Score 124.5; DB 16; Length 2482;  
 Best Local Similarity 26.2%; Pred. No. 0.23; Mismatches 83; Indels 43; Gaps 8;  
 Matches 56; Conservative 32; Mismatches 83; Indels 43; Gaps 8;  
 Qy 77 LAGNAYLQK-----TANHLYQDLOREVGSSLKEINFLMSVLYQKEFLH----- 126  
 Db 314 lisetlslekemssisinkreisitqngtikineasinslqekmnl1qksesfanyid 373  
 Qy 127 -TSKDISAVS ---QDFYSCLOQFRDNKYKGFESELIDEXYNSTEEMRKL-----FSQEIIIA 176  
 Db 374 ereksiseisdqyqkjeeklllqrceetgnayedisqskyraeqeknskleillnsectsice 433  
 Qy 177 DLKGSYASLREBT----RFITPLAEVRLAHNQESTAAIEELKTIRDLSLREI----- 227  
 Db 434 nrkneiqkleafakehqefitklatfaeraer--nqnlml---eletvqdalrsmtndq 486  
 Qy 228 -----GQLSOLSKTITSQIALQRELESSDCSQ 254  
 Db 487 mnsksaeagg1kqeimt1keeqnkmpqevndlq 520

RESULT 7  
 ID AAW23996 standard; Protein; 2482 AA.  
 XX  
 AC AAW23996;  
 DT 28-MAY-1998 (first entry)  
 DE Human mitosin amino acid sequence.  
 XX  
 KW Mitosin; Phosphoprotein; mitotic cell cycle; antibody; analogue;  
 KW inhibition; M phase; Antagonist; hyperproliferative cell; cancer;  
 KW leukaemia; lymphoma; chromosome segregation.  
 OS Homo sapiens.

FH	Key	Location/Qualifiers	QY	127	-TSKDLSAVS----QDFYSCLOGFQRDNYKGFFESLLEDKNSTEEMKL-----FSQEILIA 176
FT	Domain	/note= "leucine heptad repeat"	Db	374	erelksiseisdsqykqekillqrcegnayedlsqykaaqeknskicllnecslce 433
FT	Domain	564..362	QY	177	DLKGSVASLREEI----RFLPPLAEEVRLAHNQESSATAEELKTIIRDLSRDEI---- 227
FT	Domain	564..593	Db	434	nrkneleqikeafakehgefktlfaeerr--nqlml---elelvqqairsemtdq 486
FT	Domain	1387..1413	QY	228	-----SQLSQSKTUTJSQIAQRKASSDLSQ 254
FT	Domain	1885..1962	Db	487	nnsksseaggllkgemtikkeeqnkmqkevdnlqe 520
FT	Domain	2165..2188	QY	228	-----SQLSQSKTUTJSQIAQRKASSDLSQ 254
FT	Misc-difference	2188	Db	487	nnsksseaggllkgemtikkeeqnkmqkevdnlqe 520
FT	Misc-difference	2300	FT		
FT	Misc-difference	2301	FT		
FT	Misc-difference	2303	FT		
FT	Misc-difference	2304	FT		
FT	Misc-difference	2305	FT		
FT	Misc-difference	2306	FT		
FT	Misc-difference	2307	FT		
FT	Misc-difference	2308	FT		
FT	Misc-difference	2309	FT		
FT	Misc-difference	2310	FT		
FT	Misc-difference	2311	FT		
FT	Misc-difference	2312	FT		
FT	Misc-difference	2313	FT		
FT	Misc-difference	2314	FT		
FT	Misc-difference	2315	FT		
FT	Misc-difference	2316	FT		
FT	Misc-difference	2317	FT		
FT	Misc-difference	2318	FT		
FT	Misc-difference	2319	FT		
FT	Misc-difference	2320	FT		
FT	Misc-difference	2321	FT		
FT	Misc-difference	2322	FT		
FT	Misc-difference	2323	FT		
FT	Misc-difference	2324	FT		
FT	Misc-difference	2325	FT		
FT	Misc-difference	2326	FT		
FT	Misc-difference	2327	FT		
FT	Misc-difference	2328	FT		
FT	Misc-difference	2329	FT		
FT	Misc-difference	2330	FT		
FT	Misc-difference	2331	FT		
FT	Misc-difference	2332	FT		
FT	Misc-difference	2333	FT		
FT	Misc-difference	2334	FT		
FT	Misc-difference	2335	FT		
FT	Misc-difference	2336	FT		
FT	Misc-difference	2337	FT		
FT	Misc-difference	2338	FT		
FT	Misc-difference	2339	FT		
FT	Misc-difference	2340	FT		
FT	Misc-difference	2341	FT		
FT	Misc-difference	2342	FT		
FT	Misc-difference	2343	FT		
FT	Misc-difference	2344	FT		
FT	Misc-difference	2345	FT		
FT	Misc-difference	2346	FT		
FT	Misc-difference	2347	FT		
FT	Misc-difference	2348	FT		
FT	Misc-difference	2349	FT		
FT	Misc-difference	2350	FT		
FT	Misc-difference	2351	FT		
FT	Misc-difference	2352	FT		
FT	Misc-difference	2353	FT		
FT	Misc-difference	2354	FT		
FT	Misc-difference	2355	FT		
FT	Misc-difference	2356	FT		
FT	Misc-difference	2357	FT		
FT	Misc-difference	2358	FT		
FT	Misc-difference	2359	FT		
FT	Misc-difference	2360	FT		
FT	Misc-difference	2361	FT		
FT	Misc-difference	2362	FT		
FT	Misc-difference	2363	FT		
FT	Misc-difference	2364	FT		
FT	Misc-difference	2365	FT		
FT	Misc-difference	2366	FT		
FT	Misc-difference	2367	FT		
FT	Misc-difference	2368	FT		
FT	Misc-difference	2369	FT		
FT	Misc-difference	2370	FT		
FT	Misc-difference	2371	FT		
FT	Misc-difference	2372	FT		
FT	Misc-difference	2373	FT		
FT	Misc-difference	2374	FT		
FT	Misc-difference	2375	FT		
FT	Misc-difference	2376	FT		
FT	Misc-difference	2377	FT		
FT	Misc-difference	2378	FT		
FT	Misc-difference	2379	FT		
FT	Misc-difference	2380	FT		
FT	Misc-difference	2381	FT		
FT	Misc-difference	2382	FT		
FT	Misc-difference	2383	FT		
FT	Misc-difference	2384	FT		
FT	Misc-difference	2385	FT		
FT	Misc-difference	2386	FT		
FT	Misc-difference	2387	FT		
FT	Misc-difference	2388	FT		
FT	Misc-difference	2389	FT		
FT	Misc-difference	2390	FT		
FT	Misc-difference	2391	FT		
FT	Misc-difference	2392	FT		
FT	Misc-difference	2393	FT		
FT	Misc-difference	2394	FT		
FT	Misc-difference	2395	FT		
FT	Misc-difference	2396	FT		
FT	Misc-difference	2397	FT		
FT	Misc-difference	2398	FT		
FT	Misc-difference	2399	FT		
FT	Misc-difference	2400	FT		
FT	Misc-difference	2401	FT		
FT	Misc-difference	2402	FT		
FT	Misc-difference	2403	FT		
FT	Misc-difference	2404	FT		
FT	Misc-difference	2405	FT		
FT	Misc-difference	2406	FT		
FT	Misc-difference	2407	FT		
FT	Misc-difference	2408	FT		
FT	Misc-difference	2409	FT		
FT	Misc-difference	2410	FT		
FT	Misc-difference	2411	FT		
FT	Misc-difference	2412	FT		
FT	Misc-difference	2413	FT		
FT	Misc-difference	2414	FT		
FT	Misc-difference	2415	FT		
FT	Misc-difference	2416	FT		
FT	Misc-difference	2417	FT		
FT	Misc-difference	2418	FT		
FT	Misc-difference	2419	FT		
FT	Misc-difference	2420	FT		
FT	Misc-difference	2421	FT		
FT	Misc-difference	2422	FT		
FT	Misc-difference	2423	FT		
FT	Misc-difference	2424	FT		
FT	Misc-difference	2425	FT		
FT	Misc-difference	2426	FT		
FT	Misc-difference	2427	FT		
FT	Misc-difference	2428	FT		
FT	Misc-difference	2429	FT		
FT	Misc-difference	2430	FT		
FT	Misc-difference	2431	FT		
FT	Misc-difference	2432	FT		
FT	Misc-difference	2433	FT		
FT	Misc-difference	2434	FT		
FT	Misc-difference	2435	FT		
FT	Misc-difference	2436	FT		
FT	Misc-difference	2437	FT		
FT	Misc-difference	2438	FT		
FT	Misc-difference	2439	FT		
FT	Misc-difference	2440	FT		
FT	Misc-difference	2441	FT		
FT	Misc-difference	2442	FT		
FT	Misc-difference	2443	FT		
FT	Misc-difference	2444	FT		
FT	Misc-difference	2445	FT		
FT	Misc-difference	2446	FT		
FT	Misc-difference	2447	FT		
FT	Misc-difference	2448	FT		
FT	Misc-difference	2449	FT		
FT	Misc-difference	2450	FT		
FT	Misc-difference	2451	FT		
FT	Misc-difference	2452	FT		
FT	Misc-difference	2453	FT		
FT	Misc-difference	2454	FT		
FT	Misc-difference	2455	FT		
FT	Misc-difference	2456	FT		
FT	Misc-difference	2457	FT		
FT	Misc-difference	2458	FT		
FT	Misc-difference	2459	FT		
FT	Misc-difference	2460	FT		
FT	Misc-difference	2461	FT		
FT	Misc-difference	2462	FT		
FT	Misc-difference	2463	FT		
FT	Misc-difference	2464	FT		
FT	Misc-difference	2465	FT		
FT	Misc-difference	2466	FT		
FT	Misc-difference	2467	FT		
FT	Misc-difference	2468	FT		
FT	Misc-difference	2469	FT		
FT	Misc-difference	2470	FT		
FT	Misc-difference	2471	FT		
FT	Misc-difference	2472	FT		
FT	Misc-difference	2473	FT		
FT	Misc-difference	2474	FT		
FT	Misc-difference	2475	FT		
FT	Misc-difference	2476	FT		
FT	Misc-difference	2477	FT		
FT	Misc-difference	2478	FT		
FT	Misc-difference	2479	FT		
FT	Misc-difference	2480	FT		
FT	Misc-difference	2481	FT		
FT	Misc-difference	2482	FT		
FT	Misc-difference	2483	FT		
FT	Misc-difference	2484	FT		
FT	Misc-difference	2485	FT		
FT	Misc-difference	2486	FT		
FT	Misc-difference	2487	FT		
FT	Misc-difference	2488	FT		
FT	Misc-difference	2489	FT		
FT	Misc-difference	2490	FT		
FT	Misc-difference	2491	FT		
FT	Misc-difference	2492	FT		
FT	Misc-difference	2493	FT		
FT	Misc-difference	2494	FT		
FT	Misc-difference	2495	FT		
FT	Misc-difference	2496	FT		
FT	Misc-difference	2497	FT		
FT	Misc-difference	2498	FT		
FT	Misc-difference	2499	FT		
FT	Misc-difference	2500	FT		
FT	Misc-difference	2501	FT		
FT	Misc-difference	2502	FT		
FT	Misc-difference	2503	FT		
FT	Misc-difference	2504	FT		
FT	Misc-difference	2505	FT		
FT	Misc-difference	2506	FT		
FT	Misc-difference	2507	FT		
FT	Misc-difference	2508	FT		
FT	Misc-difference	2509	FT		
FT	Misc-difference	2510	FT		
FT	Misc-difference	2511	FT		
FT	Misc-difference	2512	FT		
FT	Misc-difference	2513	FT		
FT	Misc-difference	2514	FT		
FT	Misc-difference	2515	FT		
FT	Misc-difference	2516	FT		
FT	Misc-difference	2517	FT		
FT	Misc-difference	2518	FT		
FT	Misc-difference	2519	FT		
FT	Misc-difference	2520	FT		
FT	Misc-difference	2521	FT		
FT	Misc-difference	2522	FT		
FT	Misc-difference	2523	FT		
FT	Misc-difference	2524	FT		
FT	Misc-difference	2525	FT		
FT	Misc-difference	2526	FT		
FT	Misc-difference	2527	FT		
FT	Misc-difference	2528	FT		
FT	Misc-difference	2529	FT		
FT	Misc-difference	2530	FT		
FT	Misc-difference	2531	FT		
FT	Misc-difference	2532	FT		
FT	Misc-difference	2533	FT		
FT	Misc-difference	2534	FT		
FT	Misc-difference	2535	FT		
FT	Misc-difference	2536	FT		
FT	Misc-difference	2537	FT		
FT	Misc-difference	2538	FT		
FT	Misc-difference	2539	FT		
FT	Misc-difference	2540	FT		
FT	Misc-difference	2541	FT		
FT	Misc-difference	2542	FT		
FT	Misc-difference	2543	FT		
FT	Misc-difference	2544	FT		
FT	Misc-difference	2545	FT		
FT	Misc-difference	2546	FT		
FT	Misc-difference	2547	FT		
FT	Misc-difference	2548	FT		
FT	Misc-difference	2549	FT		
FT	Misc-difference	2550	FT		
FT	Misc-difference	2551	FT		
FT	Misc-difference	2552	FT		
FT	Misc-difference	2553	FT		

CC cells. CENP-F can be used to detect autoimmune antibodies to  
 CC the protein, which may provide an early diagnosis for the onset  
 CC of various malignant diseases. Use of CENP-F as a cell cycle  
 CC marker allows the specific detection of G2 and M phase cells.  
 XX Sequence 3248 AA;

Query Match 9.3%; Score 124.5; DB 17; Length 3248;  
 Best Local Similarity 26.2%; Pred. No. 0.33; Indels 43; Gaps 8;  
 Matches 56; Conservative 32; Mismatches 83; Sequence 3248 AA;

```

  Qy  77 LAGNALYIQLK-----TANILHYQDLORENGSLKEINFMISVLQKEFLHISKERAT--- 126
  Db  946 llsetsliekemssislnkrieelqtqekineasngemknliqestanyid 1005
  Qy  127 -TSKDLSSAVS---QDFYSCLQFRDYKGFFSLDEYKNSTEEMRKL----FSEIIIA 176
  Db  1006 erksiseisdyykqeklillirceetgnayedlsqkyaqeknsklecllnectslic 1065
  Qy  177 DLKGSVASLREEL---RFLTPLAEEFVRLAHNOESLTATAEELKTIRDSLDET--- 227
  Db  1066 nrnneledkeafakehgiefliklaaer--nqnml---eletvqqalrsemtndq 1158
  Qy  228 -----GQLSQLSKTUTSQTAQLQRKESSDLCSQ 254
  Db  1119 nnskseadgqkgeamtikkeeqhnqmkevndlqie 1152

```

RESULT 9  
 ID ABB61012 standard; Protein; 1931 AA.  
 XX ABB61012;  
 AC ABB61012;  
 DT 26-MAR-2002 (first entry)  
 XX Drosophila melanogaster polypeptide SEQ ID NO 9828.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX Drosophila melanogaster.  
 OS Drosophila melanogaster.  
 PN WO2001171042-A2.  
 XX WO200130817-A1.

XX 27-SEP-2001 .  
 XX 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL05115.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 PS Disclosure; SEQ ID NO 9828; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL01840-ABL01675), expressed DNA  
 CC sequences (ABL01840-ABL01675) and the encoded proteins

CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at fip.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 1931 AA;

Query Match 9.0%; Score 120; DB 22; Length 1931;  
 Best Local Similarity 24.1%; Pred. No. 0.4; Mismatches 89; Indels 32; Gaps 7;  
 Matches 49; Conservative 33; Mismatches 89; Sequence 1931 AA;

```

  Qy  90 LHLYODL---QREVGSLKEINFMISVLQKEFLHISKERAT---SDLSAVSDFVSCLOCF 145
  Db  1413 lhaydkcvseyrlkgcidsnkisenlqkverhaeqqlalqgqisgrdselkqlrse1 1472
  Qy  146 RDNYK-----GEFSLLDEYKNSTEEMRKLQEIIADLKGSVASLR-----EEI 189
  Db  1473 kdaidentktyreaktglenlskavgenmsqeqfqfkq-iadiqgsvedlqikls1qev 1531
  Qy  190 R-FITPLAEEFVRLAHNQESTATAEELKTIRDSLDEIGOLSLSKLTISQUALQRKIE- 247
  Db  1532 rdhlesrneelkrikldaqelqnmvdqkqkinkssredfaklegtkldieeqirakkve1 1591
  Qy  248 ---SSDL-----CSQIRETLSS 261
  Db  1592 drrskelgevtkdcenirsdea 1614

```

RESULT 10  
 AAE02245  
 ID AAE02245 standard; Protein; 875 AA.  
 XX AAE02245;  
 AC AAE02245;  
 DT 31-JUL-2001 (first entry)  
 XX DE Domestic mite Bt11 allergen.  
 XX KW Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;  
 KW Immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;  
 KW asthma; antiallergic; antiinflammatory; immunosuppressive.  
 XX OS Blomia tropicalis.  
 XX PN WO200130817-A1.  
 XX 03-MAY-2001.  
 XX PD 03-MAY-2001.  
 XX PF 10-OCT-2000; 2000WO-AU01227.

XX PR 26-OCT-1999; 99SG-0005313.  
 PR 18-JUL-2000; 2000AU-0008842.  
 PR 18-JUL-2000; 2000AU-0008844.  
 PR 18-JUL-2000; 2000AU-0008845.  
 XX PA (UYS1-) UNIV SINGAPORE NAT.  
 XX PI Chua KY, Cheong N, Lee BW;  
 XX DR WPI; 2001-308609/32.  
 DR N-PSDB; AAD06237.

XX Novel immunogenic protein derived from house mite, Blomia tropicalis  
 PT useful for treating and diagnosing conditions involving induction of  
 PT rhinitis -  
 XX Disclosure; Page 162-166; 230pp; English.  
 PT The present invention relates to immunogenic proteins, referred to as Bt  
 PT allergens, derived from domestic mite (Blomia tropicalis). The specific  
 XX CC Bt allergens of the invention includes Bt11, Bt5 and BtA2. The  
 CC

CC Immunogenic protein is useful for preventing, reducing or ameliorating  
 CC Blomia tropicalis hypersensitivity condition such as atopic dermatitis,  
 CC immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or  
 CC asthma and for modulating an immune response directed to Bt allergen in  
 CC a subject. The Bt allergens are also useful for detecting antibody  
 CC directed to all or a part of Bt allergen in biological sample from a  
 CC subject. Antibodies to Bt allergens are also used as therapeutic or  
 CC diagnostic agents, to screen Bt immunosassays and as antagonists to  
 CC inhibit Bt activity under circumstances where temporary hypersensitivity  
 CC inhibition is required. The present sequence is BtII allergen.  
 XX SQ Sequence 875 AA;

Query Match 8.7%; Score 116; DB 22; Length 875;  
 Best Local Similarity 26.9%; Pred. No. 0.3;  
 Matches 46; Conservative 31; Mismatches 72; Indels 22; Gaps 5;  
 Qy 94 QDLOREYVSLKEINFMLSVLQKEFHLHSKDFEATTSKDLSAVSQDFYSCLOGFRDNKKGFE 153  
 Db 364 qalekrsqlekinldksklevsmllqq--tqdklrvkiad---lqklqhey--e 413  
 [REDACTED] 154 SLIDEYKNSTEEMRKLFQSEITIADLKGSVASLREEIRFLPLAAEVYRLAHNQESLTAAI 213  
 Db 414 kirdqkealarenkkkad----dlaeksgsindahrrihegeielkrlenerelaay 468  
 Qy 214 EELKTIDRSLRDEIGOLSQLSKTLTSQLQARKESSDLCSQIRETLSSPRK 264  
 Db 469 keaatirk-----geaknqrtaelaqrtdyekrlaqkeeetealrk 512  
 RESULT 11  
 AAE02242 standard; Protein: 878 AA.

XX AAE02242;

XX DT 31-JUL-2001 (first entry)

XX DE Domestic mite BtII allergen #7.

XX KW Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;

XX KW immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;

XX KW asthma; antiallergic; antiinflammatory; immunosuppressive.

OS Blomia tropicalis.

PN WO200130817-A1.

XX DD 03-MAY-2001.

XX PR 10-OCT-2000; 2000WO-AU01227.

XX PR 26-OCT-1999; 99SG-0005313.

XX PR 18-JUL-2000; 2000AU-0008842.

XX PR 18-JUL-2000; 2000AU-0008844.

XX PR 18-JUL-2000; 2000AU-0008845.

XX PA (UIYSI-) UNIV SINGAPORE NAT.

XX PI Chua KY, Cheong N, Lee BW;

XX DR WPI; 2001-308609/32.

DR N-PSDB; AAD05236.

XX PS Claim 4; Fig 3; 230pp; English.

XX CC The present invention relates to immunogenic proteins, referred as Bt

CC allergen, is derived from domestic mite, *Blomia tropicalis*. The specific  
 CC Bt allergens of the invention includes BtII, Bt5 and Bt62. The specific  
 CC immunogenic protein is useful for preventing, reducing or ameliorating  
 CC Bmonia tropicalis hypersensitivity condition such as atopic dermatitis,  
 CC immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or  
 CC asthma and for modulating an immune response directed to Bt allergen in  
 CC a subject. The Bt allergens are also useful for detecting antibody  
 CC directed to all or a part of Bt allergen in biological sample from a  
 CC subject. Antibodies to Bt allergens are also used as therapeutic or  
 CC diagnostic agents, to screen Bt immunosassays and as antagonists to  
 CC inhibit Bt activity under circumstances where temporary hypersensitivity  
 CC inhibition is required. The present sequence is BtII allergen.  
 XX SQ Sequence 878 AA;

Query Match 8.7%; Score 116; DB 22; Length 878;  
 Best Local Similarity 26.9%; Pred. No. 0.31; Mismatches 31; Indels 22; Gaps 5;  
 Matches 46; Conservative 31; Mismatches 72; Indels 22; Gaps 5;  
 Qy 94 QDLOREYVSLKEINFMLSVLQKEFHLHSKDFEATTSKDLSAVSQDFYSCLOGFRDNKKGFE 153  
 Db 367 qalekrsqlekinldksklevsmllqq--tqdklrvkiad---lqklqhey--e 416  
 Qy 154 SLIDEYKNSTEEMRKLFQSEITIADLKGSVASLREEIRFLPLAAEVYRLAHNQESLTAAI 213  
 Db 417 kirdqkealarenkkkad----dlaeksgsindahrrihegeielkrlenerelaay 471  
 Qy 214 EELKTIDRSLRDEIGOLSQLSKTLTSQLQARKESSDLCSQIRETLSSPRK 264  
 Db 472 keaatirk-----geaknqrtaelaqrtdyekrlaqkeeetealrk 515  
 RESULT 12  
 ID AAE02246  
 ID AAE02246 standard; Protein: 1017 AA.  
 XX AC AAE02246;  
 XX AC AAE02246;  
 XX AC AAE02246;  
 XX DT 31-JUL-2001 (first entry)  
 DE Domestic mite BtII allergen polymorphic variant.  
 KW Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;  
 KW immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;  
 KW asthma; antiallergic; antiinflammatory; immunosuppressive.  
 OS Blomia tropicalis.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 41  
 FT Misc-difference 42 /note= "Encoded by TAG"  
 FT Misc-difference 56 /note= "Encoded by TAG"  
 FT Misc-difference 71 /note= "Encoded by TGA"  
 FT Misc-difference 76 /note= "Encoded by TAA"  
 FT Misc-difference 80 /note= "Encoded by TAG"  
 FT Misc-difference 86 /note= "Encoded by TGA"  
 FT Misc-difference 965 /note= "Encoded by TAA"  
 FT Misc-difference 98 /note= "Encoded by TAA"  
 FT Misc-difference 998 /note= "Encoded by TAA"  
 XX PN WO200130817-A1.  
 XX PD 03-MAY-2001.  
 XX CC

PF 10-OCT-2000; 2000WO-AU01227.  
 XX PR 26-OCT-1999; 99SG-0005313.  
 PR 18-JUL-2000; 2000AU-0008842.  
 PR 18-JUL-2000; 2000AU-0008844.  
 PR 18-JUL-2000; 2000AU-0008845.  
 XX PA (UYI-) UNIV SINGAPORE NAT.  
 PI Chua KY, Cheong N, Lee BW;  
 XX DR WPI; 2001-308609/32.  
 DR N-PDB; AAD06245.  
 XX PT Novel immunogenic protein derived from house mite, Blomia tropicalis  
 PT useful for treating and diagnosing conditions involving induction of  
 PT immunoresponse to mite, such as allergic asthma, atopic dermatitis,  
 PT rhinitis -

**Claim 6:** Fig 7; 230PP; English.

The present invention relates to immunogenic proteins, referred as Bt alleagen, is derived from domestic mite, Blomia tropicalis. The specific Bt allergens of the invention includes Bt11, Bt10, Bt5 and BtA2. The immunogenic protein is useful for preventing, reducing or ameliorating Blomia tropicalis hypersensitivity condition such as atopic dermatitis, immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or asthma and for modulating an immune response directed to Bt allergen in a subject. The Bt allergens are also useful for detecting antibody directed to Bt or a part of Bt allergen in a biological sample from a subject. Antibodies to Bt allergens are also used as therapeutic or diagnostic agents, to screen Bt immunoassays and as antagonists to inhibit Bt activity under circumstances where temporary hypersensitivity inhibition is required. The present sequence is a protein encoded by Bt11 polymorphic variant.

XX Sequence 1017 AA;

Query Match 8.7%; Score 116; DB 22; Length 1017;  
 Best Local Similarity 26.9%; Pred. No. 0.37; 5;  
 Matches 46; Conservative 31; Mismatches 72; Indels 22; Caps 5;

QY 94 QDLIREVGSKLEINFMVSLOKEFLHLSKEFATTSKDLAVSQDFEYSCLQFRDNYKGFE 153  
 Db 453 qalekrksqlekinldksklevsmlieg---tqkirkvklad---19k1qhey--e 502

QY 154 SLIDEYKNSTEEMRKLSQEITADLKGSVSLREERFLPAAEVRRALAHNQESTLAAI 213  
 Db 503 kirdqkealarenkklad----dlaeakqsqindahrhieqekrknereelaay 557

QY 214 EELKTIIDSLRDEIGQLSOLSKLUTSQALQRKESSEDCSQRETLLSSPRK 264  
 Db 558 keatlk----qeeaknqrtaelaqtrdyekrlaqkeelearlk 601

Query Match 8.7%; Score 116; DB 22; Length 1177;  
 Best Local Similarity 19.7%; Pred. No. 0.45; 9;  
 Matches 48; Conservative 42; Mismatches 72; Indels 82; Gaps 9;

QY 82 LYLOQTANLHYQDILQREVGSLK-----EINFMISVLOKEFLHSKEFATTSKD 130  
 Db 673 lklrka-----leaeinsikvelrglenqgfarikmnsiekeitltdrieilse 725

QY 131 LSAVSDPFYSCLOGFDNYKGFESE-----LDEYKNSTEEMRKLSQ- 172  
 Db 726 erikse-----ledsckgiedridihkgeiaaklgkierkdklkklalenp 778

QY 173 -----ELIADLKGSVSLREI-----RFITP-----LAPEVR----- 200  
 Db 779 earevtkirevegegjklreelsresrieslnrneeiprasiseeglyvkin 838

QY 201 ---RIAHNOE---SUTAAIEBLKLTARDSLREIGQLSQTLSQIALQRKESSEDLCS 253  
 Db 839 alkantaaevlkgkjeelkateesvhskisseyrkreelektirelrekeelsk 898

QY 254 QIRE 257  
 Db 899 rmgc 902

RESULT 13  
 AAB96721 standard; Protein; 1177 AA.  
 ID AAB96721;  
 XX AC AAB96721;  
 XX DT 29-OCT-2001 (first entry)  
 DE putative P. abyssi ATPase involved in DNA repair #3.  
 XX KW Hyperthermophilic archaeon; hyperthermophilic protein.  
 XX OS Pyrococcus abyssi.  
 XX PN FR2732651-A1.  
 XX PD 27-OCT-2000.

RESULT 14  
 AAW56473 standard; Protein; 1372 AA.  
 ID AAW56473;  
 XX AC AAW56473;  
 XX DT 14-AUG-1998 (first entry)  
 DE Protein with Rho protein-combining and kinase activity.  
 XX KW Rho protein-binding activity; protein kinase activity; inhibitor;  
 KW smooth muscle fibre formation; smooth muscle contraction;  
 KW circulatory disease; treatment; tumour formation; metastasis inhibitor;  
 KW autoimmune disease; platelet aggregation inhibitor.  
 XX OS Bos sp.





GenCore version 4.5  
 Copyright (c) 1993 - 2000 Compugen Ltd.  
  
 OM protein - protein search, using sw model  
 Run on: August 13, 2002, 09:24:18 ; search time 23.55 seconds  
 (without alignments)  
 283.150 Million cell updates/  
  
 Title: US-09-673-763-14  
 Perfect score: 1333  
 Sequence: 1 MTTPTLIVIPEPPAPSYA.....QIRETFLSSPRKSASPTKSS 273  
  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5  
  
 Searched: 231628 seqs, 24425594 residues  
  
 total number of hits satisfying chosen parameters: 231628  
 Minimum DB seq length: 0  
 Maximum DB seq length: 200000000  
  
 Post processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries  
  
 Database : Issued\_Patents\_AA:  
 1: /cgcn2\_6/pptodata/2/1aa/5A\_COMB.pep: \*  
 2: /cgcn2\_6/pptodata/2/1aa/5B\_COMB.pep: \*  
 3: /cgcn2\_6/pptodata/2/1aa/6A\_COMB.pep: \*  
 4: /cgcn2\_6/pptodata/2/1aa/6B\_COMB.pep: \*  
 5: /cgcn2\_6/pptodata/2/1aa/7A\_COMB.pep: \*  
 6: /cgcn2\_6/pptodata/2/1aa/7B\_COMB.pep: \*  
 7: /cgcn2\_6/pptodata/2/1aa/8A\_COMB.pep: \*  
 8: /cgcn2\_6/pptodata/2/1aa/8B\_COMB.pep: \*  
 9: /cgcn2\_6/pptodata/2/1aa/9A\_COMB.pep: \*  
 10: /cgcn2\_6/pptodata/2/1aa/9B\_COMB.pep: \*

103	228	776	1	US-07-603-133B-17
103	229	776	1	US-07-603-133B-20
103	50	776	3	US-08-397A-15
102.5	331	775	1	US-07-603-133B-14
101	7.6	775	1	US-07-603-133B-12
100.5	13	756	4	US-09-005-198-9
98.5	7.4	683	6	5210183-3
98.5	7.4	1147	3	US-08-470-160-5
98.5	7.4	1147	3	US-08-466-491-5
98.5	7.4	1147	4	US-08-466-662-5
98.5	8	3289	2	US-08-477-451-2,
98	7.4	258	4	US-09-227-357-18
98	7.4	337	1	US-08-317-223-3
98	7.4	337	1	US-08-445-135-4
98	7.4	337	3	US-09-059-84A-3
98	3	337	3	US-08-213-632-3
98	7.4	337	5	PCT-US95-12675-3
98	7.4	1248	2	US-09-027-211-1
95	5	1248	2	US-09-027-211-1

ATTACHMENT

```

Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/podata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/podata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/podata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/podata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/podata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/podata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of +b result b.

```

```

Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/pctodata/2/1
2: /cgn2_6/pctodata/2/1
3: /cgn2_6/pctodata/2/1
4: /cgn2_6/pctodata/2/1
5: /cgn2_6/pctodata/2/1
6: /cgn2_6/pctodata/2/1

Pred. NO. is the number of result
score greater than or equal to the

```

SUMMARIES						
Result No.	Score	Query Match	Length	DB ID	Description	
1	124.5	9.3	2482	1	US-08-328-254-6	Sequence 6, App1
2	124.5	9.3	3248	1	US-08-353-709-1	Sequence 1, App1
3	124.5	9.3	3248	5	PCT-US95-1162116-1	Sequence 1, App1
4	117	8.8	775	1	US-07-603-113B-16	Sequence 16, App1
5	115.5	8.7	1388	2	US-08-685-5176-1	Sequence 1, App1
6	113.5	8.5	775	1	US-07-603-113B-15	Sequence 15, App1
7	111.5	8.4	1388	2	US-08-385-5576-4	Sequence 4, App1
8	107.5	8.1	546	2	US-09-067-331-1	Sequence 1, App1
9	107.5	8.1	546	4	US-09-360-490-1	Sequence 1, App1
10	106	8.0	477	1	US-08-402-217A-3	Sequence 1, App1
11	106	8.0	477	1	US-08-700-178-3	Sequence 3, App1
12	106	8.0	477	1	US-08-995-654-3	Sequence 3, App1
13	106	8.0	604	2	US-08-468-377B-12	Sequence 12, App1
14	106	8.0	604	2	US-08-468-579B-12	Sequence 12, App1
15	106	8.0	604	3	US-08-468-577B-12	Sequence 12, App1
16	106	8.0	2101	1	US-08-685-390-4	Sequence 4, App1
17	106	8.0	2101	1	US-08-470-500-4	Sequence 4, App1
18	106	8.0	2101	1	US-08-467-781-4	Sequence 4, App1
19	106	8.0	2101	1	US-08-195-487-4	Sequence 4, App1
20	106	8.0	2101	2	US-08-483-224-4	Sequence 4, App1
21	106	8.0	2101	4	US-09-452-291-1	Sequence 1, App1
22	106	8.0	2101	5	PCT-US93-05160-4	Sequence 4, App1
23	104	7.8	1354	3	US-08-685-871-2	Sequence 2, App1
24	103.5	7.8	712	2	US-08-468-576B-17	Sequence 17, App1
25	103.5	7.8	712	2	US-08-468-779B-17	Sequence 17, App1
26	103.5	7.8	712	3	US-08-468-577B-17	Sequence 17, App1
27	103	7.7	747	3	US-08-080-397A-17	Sequence 17, App1

Db 314 LLSETLSLEKKEMSSIIISLNKREIEBLQENTLKEINASLNQEKMNLIQKSESFANY ID 373  
 Qy 127 -TSKDLSAVS---QDFYSCLOGFRDNYKFFESLDEYKSTEENRKL---FSQEITLA 176  
 Db 374 EREKSISELSDQYKQELLQLRCETGNAEDLSQKYKAQEKNASKLECLNNECTSLCE 433  
 Qy 177 DLKGSVASLREEI---RFLTPLAEEVRLAHNQESTTAIEBLKLTIRDSLDEI--- 227  
 Db 434 NRKNELQLEKFAKEHQLTFLKAAFER---NONMLM---ELETVQQLRSEMTDNO 486  
 Qy 228 -----GQLSQLSKLTLSQALQRKESSTDCSQ 254  
 Db 487 NNSKSEAGGLQKETMLKEQNMQKEVNLLQE 520

RESULT 2  
 Sequence 1, Application US/08353700  
 GENERAL INFORMATION:  
 APPLICANT: YEN, TIMOTHY J.  
 ADDRESS: DANN, DORFMAN, HERREL AND SKILLMAN  
 STREET: 1601 MARKET STREET, SUITE 720  
 CITY: PHILADELPHIA  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19103-2307  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/16216-1  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/353,700  
 FILING DATE: 09-DEC-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: REED, Janet E.  
 REGISTRATION NUMBER: 36,252  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (215) 563-4100  
 TELEFAX: (215) 563-4044  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3248 amino acids  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: HUMAN  
 US-08-353-700-1

Query Match 9.3%; Score 124.5; DB 1; Length 3248;  
 Best Local Similarity 26.2%; Pred. No. 0.034; Mismatches 83; Indels 43; Gaps 8;

Db 946 LLSETLSLEKKEMSSIIISLNKREIEELTQENGTLKEINASLNQEKMNLIQKSESFANY ID 1005  
 Qy 77 LAGNALYLQK---TANLHYLQDLOREVGSLKEINEMLSVLQKEFLHSKEFAT--- 126  
 Db 946 LLSETLSLEKKEMSSIIISLNKREIEELTQENGTLKEINASLNQEKMNLIQKSESFANY ID 1005  
 Qy 127 -TSKDLSAVS---ODYSCLQFRDNYKFFESLDEYKSTEENRKL---FSQEITLA 176  
 Db 1006 EREKSISELSDQYKQELLQLRCETGNAEDLSQKYKAQEKNASKLECLNNECTSLCE 1065  
 Qy 177 DLKGSVASLREEI---RFLTPLAEEVRLAHNQESTTAIEBLKLTIRDSLDEI--- 227  
 Db 1066 NRKNELQLEKFAKEHQLTFLKAAFER -NONMLM---ELETVQQLRSEMTDNO 1118  
 Qy 228 -----GQLSQLSKLTLSQALQRKESSTDCSQ 254

Db 1119 NNSKSEAGGLQBIIMTLKEEQNMQKEVNNDLQE 1152

RESULT 4  
US-07-603-133B-16

; Sequence 16, Application US/07603133B  
; Patent No. 5298244

; GENERAL INFORMATION:

; APPLICANT: Redmond, Mark J.  
; APPLICANT: Ijaz, Mohammad K.  
; APPLICANT: Parker, Michael D.

; TITLE OF INVENTION: ASSEMBLED VIRAL PARTICLES AND THEIR  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 545 Middlefield Road, Suite 200  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA

; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/603,133B

; FILING DATE: 199010025  
; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:  
; NAME: Robins, Roberta L.  
; REGISTRATION NUMBER: 33,208

; REFERENCE/DOCKET NUMBER: 9313-0004.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 327-7250

; TELEFAX: (415) 327-2951  
; TELEX: 706441

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 775 amino acids

; TYPE: AMINO ACID

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

US-07-603-133B-16

; APPLICANT: Nakano, Takeshi  
; APPLICANT: Ito, Masaki  
; TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685,576  
; FILING DATE: 24-JUL-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-325129  
; FILING DATE: 20-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-17150  
; FILING DATE: 05-JAN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-131206  
; FILING DATE: 26-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bent, Stephen A.  
; REGISTRATION NUMBER: 904136  
; REFERENCE/DOCKET NUMBER: 16887/843  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1388 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-685-576-1

Query Match 8.7%; Score 115.5; DB 2; Length 1388;  
Best Local Similarity 21.5%; Pred. No. 0, 065; Matches 45; Conservatve 34; Mismatches 81; Indels 49; Gaps 4;  
Qy 84 LQKTANLHYLDLQRNGSLK-EINFMVLQKEPLHLSKEFATTSK-DLSAVSQDPEYSCLOGFRDNYKGF 142  
Db 667 LEKRQQLERFTDLEKKNNMEIDMYQLKVYIQSLQEETEHKATKARLA----- 716  
Qy 143 QGFDRNYKGPFESLUDYEKNSTEEMRKLFSE----- IIADLKGSV 182  
Db 717 ---DINKYTESIEPAKSEAMKEMKFLSILERTLQKVENNLLEAEKRCSSILDCLKQQ 772  
Qy 183 ASLREIRFLTPLAEEFVRRLA----- HNOEQLTANIEELKTIOSLRDETG 228  
Db 773 QKINELLKQDVLNEDVRNLTLLKIEQETOKRCLTQNDLKMQTQYNTLKMSEROLKQENN 832  
Qy 229 QLSQSLTSLTSQIAQRKQESSDLCSSQIRE 257  
Db 833 HLLEMKMSLEKONAELRKERQDADGQMK 861

RESULT 5  
US-08-685-576-1  
; Sequence 1, Application US/08685576  
; GENERAL INFORMATION:  
; APPLICANT: Kubuchi, Kozo  
; APPLICANT: Iwamatsu, Akihiro  
; GENERAL INFORMATION:  
; RESULT 6  
US-07-603-133B-15  
; Sequence 15, Application US/07603133B  
; Patent No. 5298244  
; GENERAL INFORMATION:

APPLICANT: Redmond, Mark J.  
 APPLICANT: Ijaz, Mohammed K.  
 APPLICANT: Parker, Michael D.  
 TITLE OF INVENTION: ASSEMBLED VIRAL PARTICLES AND THEIR USE IN A VACCINE TO ROTAVIRAL DISEASE  
 NUMBER OF SEQUENCES: 30  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morrison & Foerster  
 STREET: 545 Middlefield Road, Suite 200  
 CITY: Menlo Park  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94025  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US-07/603,133B  
 FILING DATE: 1990/10/25  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Robins, Roberta L.  
 REGISTRATION NUMBER: 3,208  
 REFERENCE/DOCKET NUMBER: 9313-004.00  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 327-7250  
 TELEX: (415) 327-2951  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 775 amino acids  
 TYPE: AMINO ACID  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 US-07-603-133B-15

Query Match 8.5%; Score 113.5; DB 1; Length 775;  
 Best Local Similarity 23.2%; Pred. No. 0.042; Gaps 9;  
 Matches 55; Conservative 45; Mismatches 100; Indels 37; Gaps 9;  
 Matches 43; Conservatve 40; Mismatches 91; Indels 21; Gaps 4;  
 QY 4.3 LIGTIGFLAIGHGLIAPQITIVLAFATSLAGNAYLQLKTAN-LHYODILOREVG 101  
 Db 446 LYGLPAFNPGSCHHEY-YELLAGRFSSFLI-----VPSNDYQTPTIMNSVYRDLERQG 498  
 102 SLKEINFMLSVLQKEFLHLSKSFATIS-TL-SAVSDDFYSCLGQPIRDNFKGFESELDEYK 160  
 Db 499 DIRE-----EFINSSQEIAMTQLIDALLPLDMFSM-----FSGIKSTIDAAK 541  
 QY 161 NSTTEEMRKLFSQE----IADLKGSVASLREEIRFLTPLAREVRLLAHNOHSLTAIEEL 216  
 Db 542 SMATKMKFKRSGLATTSISETTGSLNSNAASSISSRISRNISL---VWTDVSEQI 597  
 QY 217 KTIRDSSLRDEIGQLSQLSKLTISQIALQRKES--SDLCSQIRETSPRSKASAPST 270  
 Db 598 AGSSDSVSNISTQMSATSRRLREITTQTEGMNFDDISAAVLTKIDRSTHISPD 654  
 RESULT 7  
 US-08-685-576-4  
 Sequence 4, Application US/08685576  
 Patent No. 5906819  
 GENERAL INFORMATION:  
 APPLICANT: Kaiuchi, Kozo  
 APPLICANT: Iwamatsu, Akihiro  
 APPLICANT: Nakano, Takeshi  
 APPLICANT: Ito, Masaki  
 APPLICANT: Takahashi, No. 5906819uaki  
 TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE  
 NUMBER OF SEQUENCES : 16

RESULT 8  
 US-09-667-31-1  
 Sequence 1, Application US/09067351  
 Patent No. 5934081  
 GENERAL INFORMATION:  
 APPLICANT: Tang, Y. Tom  
 APPLICANT: Hillman, Jennifer L.  
 APPLICANT: Corley, Neil C.  
 APPLICANT: Baughn, Mariah  
 TITLE OF INVENTION: HUMAN KERATINS

NUMBER OF SEQUENCES: 6  
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304

COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/067,351  
 FILING DATE: US/09/067,351  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CERRONE, MICHAEL C.  
 REGISTRATION NUMBER: 39,132  
 REFERENCE/DOCKET NUMBER: PF-0511 US  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 546 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: PANCUT02  
 CLONE: 1467090

US-09-067-351-1

Query Match Similarity 8.1%; Score 107.5; DB 2; Length 546;  
 Best Local Similarity 24.1%; Pred. No. 0.091; Gaps 7;  
 Matches 48; Conservative 31; Mismatches 73; Indels 47; Gaps 7;

Qy 80 NALYLOQTANLHYQDIOREYGSLK-EINEMLSVLQKBFHLSKEFATTSKDLSAVSQDF 138  
 Db 279 DAAIMSKV-----ELEAKVDALNDINEFLRPLNTELEQSQISDTSVLS----- 325  
 Qy 139 YSCLGQFRDNYKG - FESTSLDEYKNTSTEEMRKLFSEELIA-----DL 178  
 Db 326 -----MDNSRSLDLGIIAEVKAQEEMARCSRAEAANYQTKEETLQAQKGDDL 378  
 Qy 179 KGSVASLREERFLPLAEVRRLAHNOESLTAAIEELKTIRDSLDEIGQLSQLSKTL 238  
 Db 379 RNRTRNEISEMNRNIAQRLOEIDNIKNQRAKLEATAEE----RGELALKDARAQEE 433  
 Qy 239 SQIALQRKESSESDICSQIRE 257  
 Db 434 LERALQRK-QDMARQIRE 451

RESULT 10  
 US-08-402-217A-3  
 Sequence 3, Application US/08402217A  
 Patent No. 557301

GENERAL INFORMATION:  
 APPLICANT: Hawkins, Phillip R.  
 APPLICANT: Wilde, Craig G.  
 APPLICANT: Seilhamer, Jeffrey J.  
 TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN UMBILICAL VEIN ENDOTHELIAL CELLS  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: Palo Alto  
 STATE: CA  
 ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/402,217A  
 FILING DATE: 10 MAR 1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Luther, Barbara J.  
 REGISTRATION NUMBER: 33954  
 REFERENCE/DOCKET NUMBER: PF-0028US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-855-0555  
 TELEFAX: 415-852-0195  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 477 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 IMMEDIATE SOURCE:  
 LIBRARY: mouse  
 CLONE: GI 53979  
 US-08-402-217A-3

Query Match Score 106; DB 1; Length 477;  
 Best Local Similarity 23.9%; Pred. No. 0.1;  
 Matches 43; Conservative 36; Mismatches 77; Indels 24; Gaps 7;

```

Qy  87 TANHLYQDQREVG-SIKEINFMSVLOKEELHISKEFATTSKDLAYSQDFYSCLQF 145
Db  160 TAQLESEQKYNDTAQSLRDVTAAQESTQEKYNDTAQSL 219
Qy  146 RDNYKGPFESLLDEYKST---EEMR- -KLFSOETIADLKGSVASLREEIRLPLAEEV 199
Db  220 RD---VSQLESYKSTKEIEDKLLENLTQEKVMAEKSVEDQGQILIAESTQEY 257
Qy  200 RRLAHN-QESLTAAIEKLTIRDSELDEIGOL-SQLSKTLTSQALQRRESSDLCSQIRE 257
Db  276 ARMVQDLQRNSTLKEEIKITSSFLEKLTDLKNQI-----RQQDEFERKQLEE 324

```

RESULT\_11  
 Sequence 3, Application US/08700178  
 Patent No. 5783669 5700912  
 GENERAL INFORMATION:  
 APPLICANT: Hawkins, Phillip R.  
 APPLICANT: Wilde, Craig G.  
 APPLICANT: Seilhamer, Jeffrey J.  
 TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: Windows Version 2.0  
 SOFTWARE: FastSEQ for Windows  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/995,654  
 FILING DATE: December 22, 1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/700,178  
 FILING DATE: August 20, 1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/402,217  
 FILING DATE: March 10, 1995  
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0028-2 DIV  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-855-0555  
 TELEFAX: 650-845-4166  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 477 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-995-654-3

Query Match 8.0%; Score 106; DB 3; Length 477;  
 Best Local Similarity 23.9%; Pred. No. 0.1;  
 Matches 43; Conservative 36; Mismatches 77; Indels 24; Gaps 7;

87 TANLHYTDQREV-SKEINEMLSVIQLKEEFTSKDLSAVSQDFYSCLOGF 145  
 Db 160 TAQESECKYNDTAQSLRDVPAQLESQEKYNDTQAQSLRDYTAQLESQEKYNDTAQSL 219

Qy 146 RDNYKGPFESLIDEVKNST---EEMR--KLSQEITADLKGSVASIRETRFLPLADEV 199  
 Db 220 RD---VSAGLESYKSTKLKEEDLKLNLTIQEKYAMMAERSVEDYQQQLTAESTNEY 275

Qy 200 RRLAHN-QESLTAAIEELKTDLIRDLEIGL-SOLSKLTLSQIALQRKESSDLGQSIRE 257  
 Db 276 ARMVQDLQRNSTLKEEEKITESSFLKEDLKKNOL--RQQDEDFRKQLEE 324

RESULT 13

US-08-468-576B-12

Sequence 12, Application US/08468576B  
 Patent No. 5981700

## GENERAL INFORMATION:

APPLICANT: Rabin, Daniel

TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS  
 TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING  
 NUMBER OF SEQUENCES: 19

## CORRESPONDENCE ADDRESS:

ADDRESSEES: Sprung Kramer Schaefer & Briscoe  
 STREET: 660 White Plains Road  
 CITY: Tarrytown  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10591-5144

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage

COMPUTER: Apple Macintosh  
 OPERATING SYSTEM: System 7.5  
 SOFTWARE: WordPerfect

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,576B  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/239,276  
 FILING DATE: 05-MAY-1994

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/872,646  
 FILING DATE: 08-JUN-1992

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/715,181  
 FILING DATE: 14-JUN-1991

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/441,703  
 FILING DATE: 04-DEC-1989

APPLICATION NUMBER: US 07/312,543  
 FILING DATE: 17-FEB-1989

ATTORNEY/AGENT INFORMATION:  
 NAME: Kurt G. Briscoe  
 REGISTRATION NUMBER: 33,141  
 REFERENCE/DOCKET NUMBER: MDI 251.7-KGB  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (914) 332-1700  
 TELEFAX: (914) 332-1844  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 604 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-08-468-576B-12

Query Match 8.0%; Score 106; DB 2; Length 604;  
 Best Local Similarity 25.2%; Pred. No. 0.14;  
 Matches 38; Conservative 37; Mismatches 48; Indels 28; Gaps 5;

Qy 97 QREVGSIKEINFLMSVYIQLKEEFTSKDLSAVSQDFYSCLOGF 156  
 Db 121 ERRKGSLADV---VDTLKQKKL-----EEMRTEDEDDSM-----EKLL 157

Qy 157 DEYKNSTEMRKLKFSEQTIAADLGKGSVASYLREEIRFLPLADEVVRLAHNQESTTAIEEL 216  
 Db 158 S--KWKERNERLNSELLGEIKGPESLAKEROLSTM---ITQLISLREQOLAAHDEQ 212

Qy 217 KTIRDSLDETGLQSKLTSKLTQALORKKE 247  
 Db 213 KKLAAQSIEKRQQNMDLAROOQEQTAROOQ 243

RESULT 14

US-08-468-579B-12

Sequence 12, Application US/08468579B  
 Patent No. 5981700

## GENERAL INFORMATION:

APPLICANT: Rabin, Daniel

TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS  
 TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING  
 NUMBER OF SEQUENCES: 19

## CORRESPONDENCE ADDRESS:

ADDRESSEES: Sprung Kramer Schaefer & Briscoe  
 STREET: 660 White Plains Road  
 CITY: Tarrytown  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10591-5144

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage

COMPUTER: Apple Macintosh  
 OPERATING SYSTEM: System 7.5  
 SOFTWARE: WordPerfect

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,579B  
 FILING DATE: 06-JUN-1995

## CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/239,276  
 FILING DATE: 05-JUN-1994

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/872,646  
 FILING DATE: 08-JUN-1992

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/715,181  
 FILING DATE: 14-JUN-1991

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/441,703  
 FILING DATE: 04-DEC-1989

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/312,543  
 FILING DATE: 17-FEB-1989

ATTORNEY/AGENT INFORMATION:  
; NAME: Kurt G. Briscoe  
; REGISTRATION NUMBER: 33,141  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (914) 332-1700  
; TELEFAX: (914) 332-1844  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 604 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-468-579B-12

Query Match 8.0%; Score 106; DB 2; Length 604;  
Best Local Similarity 25.2%; Pred. No. 0.14;  
Matches 38; Conservative 37; Mismatches 48; Indels 28; Gaps 5;

97 QREVGSILKEINFMVLSQLKRFFLHLSKEFATTSKDLAVSQDFYSCLOGFRDNYKGFEISLL 156  
Db 121 ERRKGSILADY --VDTLQKKL-----EKL 157

157 DEYKNSTEEMRKLFSEOIDLKGKVASLREEIRFLPLAEEVRLAHNQESTAAIEEL 216  
Qy 158 S--KDWKEKMERLNTESELLGEIKTPESLAKEKERQLSTIM--ITQLISUREQLIAADEQ 212

217 KTIRDSDLRDEIGOLSQLSKLTSQLAORKE 247  
Qy 213 KKLAASQIEKQRQMDLARQQEQIAROOQ 243  
Db 213 KKLAASQIEKQRQMDLARQQEQIAROOQ 243

RESULT 15  
US-08-468-579B-12  
; Sequence 12, Application US/08468577B  
; Patent No. 6001804  
; GENERAL INFORMATION:  
; APPLICANT: Rabin, Daniel  
; TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sprung Kramer Schaeffer & Briscoe  
; STREET: 660 White Plains Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10591-5144  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: System 7.5  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,577B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 514  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/239,276  
; FILING DATE: 05-MAY-1994  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/872,646  
; FILING DATE: 08-JUN-1992  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/715,181  
; FILING DATE: 14-JUN-1991  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/441,703  
; FILING DATE: 04-DEC-1989  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/312,543  
; FILING DATE: 17-FEB-1989  
; ATTORNEY/AGENT INFORMATION:

NAME: Kurt G. Briscoe  
; REGISTRATION NUMBER: 33,141  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (914) 332-1700  
; TELEFAX: (914) 332-1844  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 604 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-468-579B-12

Query Match 8.0%; Score 106; DB 3; Length 604;  
Best Local Similarity 25.2%; Pred. No. 0.14;  
Matches 38; Conservative 37; Mismatches 48; Indels 28; Gaps 5;

Qy 97 QREVGSILKEINFMVLSQLKRFFLHLSKEFATTSKDLAVSQDFYSCLOGFRDNYKGFEISLL 156  
Db 121 ERRKGSILADY --VDTLQKKL-----EKL 157

157 DEYKNSTEEMRKLFSEOIDLKGKVASLREEIRFLPLAEEVRLAHNQESTAAIEEL 216  
Qy 158 S--KDWKEKMERLNTESELLGEIKTPESLAKEKERQLSTIM--ITQLISUREQLIAADEQ 212

217 KTIRDSDLRDEIGOLSQLSKLTSQLAORKE 247  
Qy 213 KKLAASQIEKQRQMDLARQQEQIAROOQ 243  
Db 213 KKLAASQIEKQRQMDLARQQEQIAROOQ 243

Search completed: August 13, 2002, 09:24:21  
Job time: 152 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - Protein search, using sw model

Run on: August 13, 2002, 09:24:53 ; Search time 29.79 Seconds

(without alignments)  
880.577 Million cell updates/sec

Perfect score: US-09-673-763-14

Sequence: 1 MTTPTLIVPPSPAPSYA,.....QIRETLLSSPRKSASPSTKSS 273

Scoring table: BL00UM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0\*

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_71.\*

1: p1r1;\*

2: p1r2;\*

3: p1r3;\*

4: p1r4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1311	98.3	273	C71553	Probable inclusion membrane protein
2	714	53.6	276	H81707	Inclusion membrane
3	162.5	12.2	355	S61491	Hypothetical prote
4	154.5	11.6	390	A71108	Hypothetical prote
5	154.5	11.6	390	B86514	Intermediate filam
6	128.5	9.6	472	S41720	Hypothetical prote
7	128.5	9.6	1957	T38077	Probable chromos
8	126	9.5	1179	F71190	nuclear fusion pro
9	124	9.3	347	S51348	tein - human
10	122.5	9.2	3259	A56539	Outer layer protei
11	122	9.2	775	VPRW8	n - human
12	121.5	9.1	3225	I52300	Probable membrane
13	120.5	9.0	1286	T16507	Hypothetical prote
14	120	9.0	496	A12354	Outer layer protei
15	120	9.0	644	G82490	Chromosome segreg
16	117.5	8.8	1475	T33318	intermediate filam
17	116.5	8.7	419	G86517	serine/threonine-s
18	116.5	8.7	550	B91286	hypothetical prote
19	116.5	8.7	550	F86127	male-enhanced anti
20	116.5	8.7	824	T20477	cardiac muscle fac
21	116	8.7	776	VPRW9	Outer layer protei
22	116	8.7	1177	B75150	Chlamydia muridarum (s
23	115.5	8.7	567	T18872	inclusion membrane localised protein IncA TC0396 [Imported] - Chlamydia muridarum (s
24	115.5	8.7	1388	S70633	C;Species: Chlamydia muridarum, Chlamydia trachomatis
25	115.5	8.7	1974	T30010	C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
26	115	8.6	1325	T42722	C;Accession: H81107
27	114.5	8.6	1538	T29095	R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hick
28	114	8.6	775	VPRPG	
29	113.5	8.5	377	B90437	

30	113.5	8.5	555	C96667	unknown protein, 7 apolipoprotein A-I
31	112	8.4	391	LPRTA4	hypothetical prote
32	112	8.4	407	F72343	coiled coil protei
33	112	8.4	750	T38435	probable nuclear P
34	112	8.4	1837	T41023	dynain beta heavy outer layer protei
35	112	8.4	4588	T28667	probable vesicular
36	111.5	8.4	775	VPRW7	probable vesicular
37	111.5	8.4	1039	S62509	probable vesicular
38	111.5	8.4	1044	T50213	probable vesicular
39	111.5	8.4	1186	A1250	SMC protein essent
40	111.5	8.4	1251	A56677	neuronal cell cycl
41	111.5	8.4	1790	S67583	transport protein
42	111	8.3	432	A32936	glial fibrillary a
43	111	8.3	776	A4B731	outer capsid prote
44	110.5	8.3	533	T37253	hypothetical prote
45	110.5	8.3	2104	T38774	myosin-3 heavy cha

## ALIGNMENTS

RESULT 1	C71553	probable inclusion membrane protein A - Chlamydia trachomatis (serotype D, strain U3
C; Species:	Chlamydia trachomatis	C; Species: Chlamydia trachomatis
C; Date:	13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999	C; Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C; Accession:	C71553	C; Accession: C71553
R; Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche	R; Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche	R; Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
A; Title:	Genomic sequence of an obligate intracellular pathogen of humans: Chlamydia t	A; Title: Genomic sequence of an obligate intracellular pathogen of humans: Chlamydia t
A; Reference number:	A71570; MUID:9900809	A; Reference number: A71570; MUID:9900809
A; Status:	Preliminary	A; Status: Preliminary
A; Molecule type:	DNA	A; Molecule type: DNA
A; Residues:	1-273 <AFN>	A; Residues: 1-273 <AFN>
A; Cross-references:	GB:AE001286; GB:AE001273; PID:93328516; PID:93328516; PID:93328516	A; Cross-references: GB:AE001286; GB:AE001273; PID:93328516; PID:93328516; PID:93328516
A; Experimental source:	serotype D, strain U3/Cx	A; Experimental source: serotype D, strain U3/Cx
C; Genetics:		C; Genetics:
A; Gene:	incA	A; Gene: incA

RESULT 2

H81707

C; Species: Chlamydia muridarum

C; Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000

C; Accession: H81107

R; Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hick

C.: Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, S.							
ucleic Acids Res. 28, 1397-1406, 2000							
;title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.							
;Reference number: A81500; MUID: 20150555							
;Accession: H81707							
;Status: preliminary							
;Molecule type: DNA							
;Residues: 1-276 <TET>							
;Cross-references: GB:AE002306; GB:AE002160; NID:97190428; PIDN:AAFF9253.1; PID:9719043							
;Experimental source: strain Nigg (MoPn)							
;Genetics:							
;Gene: TC0396							
Query	Match	53.6%	Score 714;	DB 2;	Length 276;		
Best	Local Similarity	52.7%	Pred. No. 8.4e-39;				
Matches	145;	Conservative	58;	Mismatches 64;	Indels 8;	Gaps 2;	
1	MTTPTLIVIPPS-	PPAPSYS-----	ANVYPOPSLMDKIKKIAIASLILIGTIGFLA	52			
1	MTSPVLVEAPLSCYPPTHSS	TACTKRSSSLYKPSLSEIYQRVAATVSLSITGFLAI	60				
53	LGHILVGFLIAQOTIVLLALEITSLAGNALYLQKTANLHYQDLOREGVSGSKEINFMFLSV	112					
61	LGHAGFLIAQPAVLVIAFLAVLIAFLISLQNLGYLCKTAPLRLYKELQEVASLKEYVNFLKS	120					
113	LQKEFHLHKSKDLSAVSQDFSYCLQGFLRDNKGKFESLDEYKNTSTEEMRKLFQS	172					
121	VQEKFUGLSKDFATSKDLSVSDLDFHNLQDQFSSHQGFEDLLEYKNSAEDLRQFLFSQ	180					
173	EIIADLKGKSYASLREEFTPLAEEFVRLLTPLAEEFELKTIRTLDRTEGTQLFSQ	232					
181	ETVQSLKSTLKSERIKETVPLTVEEVRLAENKEDLTKVQDLQDIRKLRAEINNLSQ	240					
233	LSKTLTSQIALQKERSSDLCSQIREFLSPRSKSAS	267					
241	ASKTLSEQIASQEEENEKLYANITKALQSQYSQDSS	275					
RESULT	3						
561491	3561491	3561491	3561491	3561491	3561491	3561491	3561491
Species:	Chlamydophila psittaci	Species:	Chlamydophila psittaci	Species:	Chlamydophila psittaci	Species:	Chlamydophila psittaci
Date:	27-Apr-1996	#sequence_revision	13-Mar-1997	#text_change	31-Mar-2000	#text_change	31-Mar-2000
Accession:	S61491	Accession:	S61491	Accession:	S61491	Accession:	S61491
Author:	R.R.Rockey, D.D.; Heinzen, R.A.; Hackstadt, T.	Author:	R.R.Rockey, D.D.; Heinzen, R.A.; Hackstadt, T.	Author:	R.R.Rockey, D.D.; Heinzen, R.A.; Hackstadt, T.	Author:	R.R.Rockey, D.D.; Heinzen, R.A.; Hackstadt, T.
Journal:	Microbiol. 151, 617-626, 1995	Journal:	Microbiol. 151, 617-626, 1995	Journal:	Microbiol. 151, 617-626, 1995	Journal:	Microbiol. 151, 617-626, 1995
Title:	Cloning and characterisation of a Chlamydia psittaci gene coding for a protein	Title:	Cloning and characterisation of a Chlamydia psittaci gene coding for a protein	Title:	Cloning and characterisation of a Chlamydia psittaci gene coding for a protein	Title:	Cloning and characterisation of a Chlamydia psittaci gene coding for a protein
Reference number:	S61491; MUID: 95302975	Reference number:	S61491; MUID: 95302975	Reference number:	S61491; MUID: 95302975	Reference number:	S61491; MUID: 95302975
Accession:	S61491	Accession:	S61491	Accession:	S61491	Accession:	S61491
Status:	preliminary	Status:	preliminary	Status:	preliminary	Status:	preliminary
Molecule type:	DNA	Molecule type:	DNA	Molecule type:	DNA	Molecule type:	DNA
Residues:	1-355 <ROC>	Residues:	1-355 <ROC>	Residues:	1-355 <ROC>	Residues:	1-355 <ROC>
Cross-references:	EMBL:L35036;	Cross-references:	EMBL:L35036;	Cross-references:	EMBL:L35036;	Cross-references:	EMBL:L35036;
	NID:9516598;		NID:9516598;		NID:9516598;		NID:9516598;
	PID:9516599;		PID:9516599;		PID:9516599;		PID:9516599;
Query	Match	12.28;	Score 162.5;	DB 2;	Length 355;		
Best	Local Similarity	21.4%	Pred. No. 0.0025;				
Matches	63;	Conservative	61;	Mismatches 59;	Indels 59;	Gaps 8;	
8	VIPBSPAPSYSANR-----	-PQPSLMDKIKKIAIASLILIGTIGFLAIGHLYV	58				
32	IIPISTIEPTSSAAAVGAKTAIEPEGRSPILQRCYLKTTIAALFYVGIAALVCLYLG	91					
59	FIAPQTIVLLALFITSLA-----	-GNAYLQKTAHNLYKDQLR-----EV	100				
92	SVISPSLIMLAAMLVSFVIVITAIRDGTPSQVVR-----	-HMKQIQQQEENTRLHTAV	147				
101	GSIKEINFMLS-----YLOKEFLH-----LSKEFATSKDSLAVSQDQEYSCLOGFRDNYKGFE	153					
148	ENLKAVNVELSEQINQKQLHTRLSDFGDRLEANTGFTIALDFOLSLSEEFSVGTKE	207					
QY							





Db	469	FILLYPSNDYQTPIMNSVTVRQDLERQGDLE-----EFNSLSQEIAMTQLID	518
Qy	131	LSAVSQDFYSCLQGFRDNYKGFESLDEYKNSTEEMRKLFQSQTIADELKGSVASLREIR 190	
Db	519	LALLPLDMFSM-----PSGIKSTRDAKSMTVMKKFKK-----SGLATSISETJR	565
Qy	191	FLTPLAEEYVR---LAHNQESL---TAALBELKTIRDSRDEIGQLSLSKTL-TSQAL	243
Db	566	SLSNAASVSYRSSSSKNSISSYWTDVSEQITGGSSDVRNISTQTSAISKDRRLRETT	625
Qy	244	QRKESS--DLCSQIRETLLSPRSKASAPST	270
Db	626	QTEGMNFIDISAAYLKTIDKSTHISPDY	654
RESULT	12		
	152300	Giantin - human	
N;Alternative names:	gcp372		
C;Species:	Homo sapiens (man)		
Date:	01-Nov-1996	#sequence_revision	01-Nov-1996 #text_change 20-Aug-1999
A;Accession:	I52300		
A;Title:	Biochem. Biophys. Res. Commun. 205, 1399-1408, 1994		
A;Reference number:	95100974		
A;Status:	Preliminary; translated from GB/EMBL/DBJ		
A;Molecule type: mRNA			
A;Residues:	1-3225 <REBS>		
A;Cross references:	GB:D25542; PID:9662389; PIDN:BAA05025.1; PID:9808869		
C;Superfamily:	giantin		
Query Match	9.18;	Score 121.5; DB 2; Length 3225;	
Best Local Similarity	23.6%;	Pred. No. 15; Mismatches 35;	
Matches	53;	Conservative	
Qy	94	QDLQREVGSLKEINFMLS-----LOKEFLHLSKEFATTSKDSLAV-----S	135
Db	1550	EKLVKTEISLKSXKAESTWQEKELKQLEYEILLQSEENVSAEERIQQHYYEAVRQEK	1609
Qy	136	QDFYSCLQGRDNYKGFESLDEYKNSTEEMRKLF-----SQETIADLKGSVASLREI	189
Db	1610	QELYGKLRSSTEANKKETEKOLQEAQEMEMKMRKFQAKSKQKILIEEENDRLRAEV	1669
Qy	190	R-----FLPLAEEYVRLLHNQESLTAALEELTIRDSRDEIGQLS-QL	233
Db	1670	HPAGDTAKEMETLSSNSAMKMEELERYMEYETLSKQFOSLMSKEKDLSSEEVDLKHQI	1729
Qy	234	SKTLTSQIALQRKESSDLCSQIRE-TLSASP-----RKSASPSTK	271
Db	1730	EGNVSKQANLAEATEKHDNQTNVTEEGTQSIPGETEQDLSMSTR	1774
RESULT	13		
T16507		hypothetical protein F59A6.5 - Caenorhabditis elegans	
C;Species:	Caenorhabditis elegans		
C;Accession:	T16507		
R:Nhan, M.			
C;Description:	The sequence of C. elegans cosmid F59A6.		
A;Description number:	Z18526		
A;Status:	Preliminary; translated from GB/EMBL/DBJ		
A;Molecule type: DNA			
A;Residues:	1-1286 <NHA>		
C;Genetics:			
A;Cross references:	EMBL:U41994; NID:91123047; PID:91123050; PIDN:AAA83454.1; CESP:F59A6		
A;Introns:	35/3; 335/3; 685/3; 973/3; 1097/3		
RESULT	15		
	G82490	methyl-accepting chemotaxis protein VCA0176 [Imported] - Vibrio cholerae (strain N16	

C:Species: Vibrio cholerae  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: G82490  
 R:Heidelberg, J.-F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Rasmussen, D.; Ermolayeva, M.D.; Yamatheyan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
 I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A;Reference number: A82035; MUID:20406833  
 A;Accession: G82490  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-644 <HEI>  
 A;Cross-references: GB:AE004358; GB:AE003853; NID:99657566; PIDN:AAFF96089.1; GSPDB:GN001  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C;Genetics:  
 A;Gene: VCA0176  
 A;Map position: 2

Query Match 9.0%; Score 120; DB 2; Length 644;  
 Best Local Similarity 19.2%; Pred. No. 2.7;  
 Matches 44; Conservative 50; Mismatches 77; Indels 58; Gaps 7;

---

Qy	74	TISLAGNALYIQTANTMLHYIQLDQREVGSIKEINFMLSIVLQKEFLHLSKEFATISK	--D 130	
Db	415	ISELASIAATAETAN	----QASGNADQGRNVVNKAKEAISRLAIDIENTGKVVEQ	466
Qy	131	LSAVSODFYSC1QGFR	-----DNVKGFESLLDEVKN-----S 162	
Db	467	LASTTQEIGSILDAIRGISEQTNLALNALEAARAGDQRGFAVVADEVRNLASRTASS	526	
Qy	163	TEEMRKLFSEQLIADKGSVASL	-----REELRFPLAEEVRLAHNQES 208	
Db	527	TERIQMINQ-LQNDANAVSAMDAGKTVTHQVAASDEAVQVLMSISDRIHDISRNTQ	585	
Qy	209	LPAIEELKTIRDLSRDEIGQLSQLSKTLTSQIALQRKESSDLCQIRE	257	
Db	586	VAPATEEQSTVYHTINQNEEINAIEVTS	---TAEEELADASKSLRE 630	

Search completed: August 13, 2002, 09:24:56  
 Job time: 152 sec

Scoring table:	BLOSUM62				
Searched:	Gapop 10.0 , Gapext 0.5				
Total number of hits satisfying chosen parameters:	105224				
Minimum DB seq length: 0					
Maximum DB seq length: 2000000000					
Post-processing: Minimum Match 0%					
Database :	SwissProt_40.4*				
Scoring table:	BLASTP				
Scored by:	105224 seqs, 38719550 residues				
Post-processing:	Minimum Match 100% Maximum Match 100% Listeing first 45 summaries				
Description	8				
Result No.	Score	Query Match	Length	DB ID	Description
1	128.5	9.5	1957	1 YD86_SCHPO	Q10411 schizosacch
2	124.5	9.3	3210	1 CENF_HUMAN	P19451 homo sapien
3	122	9.2	775	1 VP4_ROTTH	P11198 human rotav
4	120	9.0	496	1 Y7J3_ANASP	Q05070 anabena sp
5	116	8.7	776	1 VP4_ROTTH	P11199 human rotav
6	115	8.6	1325	1 G160_MOUSE	P25937 mus musculu
7	114	8.6	775	1 VP4_ROTTH	P23045 porcine rot
8	112.5	8.4	622	1 LAMO_DRDME	P08928 drosophila
9	112	8.4	391	1 APA4_RAT	P02551 ratus norv
10	111.5	8.4	775	1 VP4_ROTTH	P11197 human rotav
11	111.5	8.4	1044	1 YAF7_SCHPO	Q09857 schizosach
12	111.5	8.4	1790	1 USOL_YEAST	P225396 saccharomy
13	111	8.3	432	1 GFAP_HUMAN	P14136 homo sapien
14	110.5	8.3	2104	1 MYS1_SCHPO	schizosacch
15	110	8.3	774	1 VP4_ROTTH	P11200 human rotav
16	110	8.3	978	1 RA50_AQUAE	Q67124 aquifex aeo
17	109.5	8.2	776	1 VP4_ROTTH	O098778 human rotav
18	109.5	8.2	1227	1 ALM1_SCHPO	Q9utk5 schizosach
19	109	8.2	432	1 OV71_ONCYO	P31732 onchocerca
20	109	8.2	776	1 VP4_ROTTH	Q06894 canine rota
21	109	8.2	2349	1 TPR_HUMAN	P012270 homo sapien
22	108.5	8.1	482	1 K2C8_HUMAN	P05787 homo sapien
23	108.5	8.1	1453	1 Y373_BOVIN	P9Lu23 bos taurus
24	108	8.1	504	1 AINX_MOUSE	P46660 mus musculu
25	108	8.1	804	1 MEA6_HUMAN	Q15320 homo sapien
26	108	8.1	1251	1 CYA8_HUMAN	P40145 homo sapien
27	107.5	8.1	1005	1 RA50_METHJA	Q58718 methanococc
28	107	8.0	535	1 HTR1_HALNI	P33741 halobacteri
29	107	8.0	776	1 VP4_ROTTH	Q07416 feline rota
30	107	8.0	776	1 VP4_ROTTH	Q06685 human rotav
31	106.5	8.0	468	1 K2C7_HUMAN	P08729 homo sapien
32	106.5	8.0	729	1 KAR3 YEAST	P17119 saccharomy
33	106.5	8.0	1189	1 SCII_CHICK	Q90988 gallus gallus

## ALIGNMENTS

RESULT	1	YD86_SCHPO	STANDARD;	PRT; 1957 AA.
ID	YD86_SCHPO			
AC	Q10411;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Hypothetical 222.8 kDa protein C1F3.06C in chromosome I.			
GN	SPAC1F3_06C.			
OS	Schizosaccharomyces pombe (Fission Yeast).			
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomyces.			
OC	Schizosaccharomyces; Schizosaccharomyces.			
NCBI_TAXID	4896;			
RN	[1]			
SEQUENCE FROM N.A.				
RC	STRAIN=972;			
RA	Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V., Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).			
CC	DR	Z70890; CAA94624.1; -.		
CC	KW	Hypothetical protein.		
SQ	SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;			
Query Match	9.6%	Score 128.5; DB 1; Length 1957;		
Best Local Similarity	27.1%	Pred. No. 1;		
Matches	51;	Conservative 32; Mismatches 82; Indels 23; Gaps 5;		
Qy	NLHLYQDIOREYGSLLKEINFMLSVLQKELPHLSKEFATTSKDLSAVSODFYSCLGQPRDN 148			
Db	345 NEKLEKLLENTIGLKDSTSNTSNSOLEEVAKESNRTHSQT---DAEKKLSSSEQE 400			
Qy	YKGFESELDDLEYKNSTEMRKLFSEOTIADLKGSVASLREEIRFLPJAEEVRLAH---N 205			
Db	401 NKSUKGSDIDEQNLLSKDKMYKQ-----VSSQLEEAR-SSLAHTGKLAENSE 449			
Qy	206 QESTTAIEELTIRDLSRDEQQLSLSQSKLTSQLQRKESSDLQSQIRETLSSPRKS 265			
Db	450 RDQNKIKDFERQDLRACLNSSNELKERSALIDKKDQELNLRQIKE----QQK 504			
Qy	266 ASPSTKSS 273			
Db	505 VSESTQSS 512			
RESULT	2	CENF_HUMAN		

ID	CENF_HUMAN	STANDARD;	PRT;	3210 AA.	
DT	P144; Q13246; Q13171;				
DT	01-FEB-1996 (Rel. 33, Created)				
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	01-FEB-2001 (Rel. 40, Last annotation update)				
DE	CENF_F kinetochore protein (Centromere protein F) (Mitosin) (AH antigen).				
DE	DE				
GN	Homo sapiens (Human).				
OS	Mammalia: Eutheria; Primates; Catarhini; Hominidae; Homo.				
OC					
NCBI_TaxID	9606;				
[1]					
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Breast carcinoma;				
RA	Liao H., Winkfein R.J., Mack G., Rattner J.B., Yen T.J.;				
RA	"CENF-F is a protein of the nuclear matrix that assembles onto kinetochores at late G2 and is rapidly degraded after mitosis.";				
RT	J. Cell Biol. 130:507-518(1995).				
[2]					
RP	SEQUENCE FROM N.A.				
RX	Medline=95379888; PubMed=7651420;				
RA	Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B., Jones D., Yang Feng T.L., Lee W.-H.; Characterization of a novel 350-kDa nuclear phosphoprotein that is specifically involved in mitotic phase progression.;				
RT	Mol. Cell. Biol. 15:5017-5029(1995).				
RL					
RN	SEQUENCE OF 2194-3210 FROM N.A.				
RX	Medline=95336444; PubMed=7612011;				
RA	Li Q., Ke Y., Hupp J.A., Fertig N., Medsger T.A. Jr., Joshi H.C.;				
RA	"A novel cell-cycle dependent 350-kDa nuclear protein: C-terminal domain is sufficient for nuclear localization";				
RT	Biochem. Biophys. Res. Commun. 212:220-228(1995).				
RL					
RN	CHARACTERIZATION.				
RP	Medline=95370296; PubMed=7642639;				
RA	Chang G.K.T., Schaar B.T., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;				
RA	"The C terminus of mitosin is essential for its nuclear localization, centromere/kinetochore targeting, and dimerization.";				
RT	J. Biol. Chem. 270:19545-19550(1995).				
[5]					
RN	CHARACTERIZATION.				
RP	Medline=98437347; PubMed=9763420;				
RA	Chang G.K.T., Schaar B.T., Yen T.J.;				
RT	"Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and HBUBRI.";				
CC	-!- FUNCTION: PROBABLY REQUIRED FOR KINETOCHORE FUNCTION, INVOLVED IN CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH RETINOBLASTOMA PROTEIN (RB), CENP-E AND BUBRI.				
CC	-!- SUBUNIT: HOMO- OR HETERO-DIMER.				
CC	-!- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS), REORGANIZATION TO THE KINETOCHORE/CENTROME (CORONAL SURFACE OF THE OUTER PLATE) AND THE SPINDLE DURING MITOSIS.				
CC	-!- DEVELOPMENTAL STAGE: GRADUALLY ACCUMULATES DURING THE CELL CYCLE.				
CC	-!- PTM: HYPERPHOSPHORYLATED DURING MITOSIS.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).				
CC	DR EMBL; U19769; AAA82889_1; -;				
DR	EMBL; U30812; AAA82835_1; -;				
DR	EMBL; U25725; AAA82838_1; -;				
DR	HSSP; P02619; ILE4 .				
DR	MM: 600236; -;				
DR	Chromosomal protein; Nuclear protein; Centromere; Coiled coil;				
RN	SEQUENCE FROM N.A.				
RX	Medline=86313706; PubMed=2839714;				
RA	Gorziglia M., Green R., Nishikawa K., Taniguchi K., Jones R., Rapikian A.Z., Chanock R.M.;				
RA	"Sequence of the fourth gene of human rotaviruses recovered from asymptomatic or symptomatic infections.";				
RT	J. Virol. 62:2978-2984(1988).				
RN	SEQUENCE OF 1-280.				
RX	Medline=86313706; PubMed=301874;				
RA	Gorziglia M., Rosino Y., Buckler-White A., Blumentals I., Glass R., Flores J., Kapikian A.Z., Chanock R.M.;				
RA	"Conservation of amino acid sequence of VP8 and cleavage region of				
RT					

RT 84-kDa outer capsid protein among rotaviruses recovered from asymptomatic neonatal infection.  
 RL Proc. Natl. Acad. Sci. U.S.A. 83: 7031-7033 (1986).  
 CC -1- SUBCELLULAR LOCATION: Outer capsid.  
 CC -1- PIM: VP8 IS ONE OF TWO TRYPSIN CLEAVAGE PRODUCTS OF VP4; THE OTHER PRODUCT IS VP5.  
 CC -1- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.  
 DR PIR: F28839; VPXW8.  
 DR PIR: F25904; VPXK16.  
 DR InterPro: IPR000416; cap\_VP4.  
 DR Pfam: PF00436; VP4; 1.  
 KW Coat protein; Glycoprotein.  
 FT CHAIN 1 775 OUTER CAPSID PROTEIN VP4.  
 FT CHAIN 1 240 OUTER CAPSID PROTEIN VP8.  
 FT CARBOHYD 247 775 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 32 56 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 583 583 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 775 AA: 87697 MW: E229A3G3D36C8018F CRC64;

Query Match 9.2%; Score 122; DB 1; Length 775;

Best Local Similarity 26.8%; Pred. No. 0.32; Mismatches 33; Indels 34; Gaps 9;

Qy 73 FITSLAGNALYLOKTA-NLHYQDLOQREVGSKKEINFLMSVLOKEPHLSEKFATSK-D 130

Db 469 FILLVPSDDYDQTPIIMSVTVRQDLERQLQDLE-----EFNSLSQEIAMTQLID 518

Qy 131 LSAYSQDFYSCLOGFRNKGPFESLSDYKUSTEENRKLFSEQELIADLGVSASLREEIR 190

Db 519 LALLPLDMFSM-----FSGKTKSTDAKSMAHMVKKK-----SCLATSISELTR 565

Qy 191 FLTFLPLAEVRR--LAHNQESL--TAAIEELKLTDRLDEIGOLSQLSSTKL-TSQLAL 243

Db 566 SLSNAASSSVRSSSIKSNISSVWTDSRQITGSSDSVRNISIQTSAAKSKRURLREIT 625

Qy 244 QRKESS--DLCSQIRPLTSSPRKSASAPST 270

Db 626 QTEGMNFIDISAALVLTKIDKSTHISPDFT 654

Query Match 9.0%; Score 120; DB 1; Length 496;

Best Local Similarity 21.9%; Pred. No. 0.73; Mismatches 54; Indels 62; Gaps 11;

Qy 38 IASLILIGTIGFLALGLHLVG-----FLAPQTIVTLIALFTSLAGNALYLOKTT 87

Db 8 IATIILIG- GVIATVGDITGTRYKARSLFLNRPKTVAVL---VYLITGLVLASATTI 61

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 176 ADLKGSVAVSLREEIRFLPLAEV----RLAHNOQSLSLTAAIEELKTIRDLSLRDEIGQL 230

Db 177 KALQGAYEVQLKTERRAYAEAKKAIEQDRELANRQIAIEQDRELANRQOALQQRDQKI 236

Qy 231 SQLSKLTS--QIALORKE 247

Db 237 SQLDKIIQNRNLTEIA-QREE 255

RESULT 4

Y7J3\_ANASP

STANDARD;

PRT;

496 AA.

RESULT 5

VP4\_ROTHN

STANDARD;

PRT;

776 AA.

ID VP4\_ROTHN

AC P1199;

DT 01-JUL-1989 (Rel. 11. Created)

DT 01-NOV-1990 (Rel. 16. Last sequence update)

DT 16-OCT-2001 (Rel. 40. Last annotation update)

DE Outer capsid protein VP4 (Hemagglutinin) (Outer layer protein VP4)

DE [Contains: Outer capsid proteins VP5 and VP6]

GN S4.

OS Human rotavirus (serotype 3 / strain MCN13).

OC dsRNA viruses; Reoviridae; Rotaviridae.

OX NCBI\_TaxID=1055;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-931308081; PubMed=8391534;

RA Wei T.-F., Ramasubramanian T.S., Pu F., Golden J.W.,

RT "Anabaena sp. strain PCC 7120 bifa gene encoding a sequence-specific

RT DNA-binding protein cloned by in vivo transcriptional interference

RT selection";

RL J. Bacteriol. 175:4025-4035(1993).

CC This SWISS-PROT entry is copy right. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC used by non profit institutions. There are no restrictions on its

CC modification and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).

CC send an email to license@isb-sib.ch).

CC

DR AP003596; BAB76692\_1;

DR EMBL; LL0036; AA004182\_1;

KW Hypothetical protein; Transmembrane; Signal; Complete proteome.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 436 HYPOTHETICAL PROTEIN ALR4393.

FT TRANSMEM 45 67 POTENTIAL.

SQ SEQUENCE 496 AA; 56375 MW;

454D67CF420AFB9E CRC64;

Qy 191 FLTFLPLAEVRR--LAHNQESL--TAAIEELKLTDRLDEIGOLSQLSSTKL-TSQLAL 243

Db 566 SLSNAASSSVRSSSIKSNISSVWTDSRQITGSSDSVRNISIQTSAAKSKRURLREIT 625

Qy 244 QRKESS--DLCSQIRPLTSSPRKSASAPST 270

Db 626 QTEGMNFIDISAALVLTKIDKSTHISPDFT 654

[2] SEQUENCE OF 1-166 FROM N.A.

RP MEDLINE-931308081; PubMed=8391534;

RA Wei T.-F., Ramasubramanian T.S., Pu F., Golden J.W.,

RT "Anabaena sp. strain PCC 7120 bifa gene encoding a sequence-specific

RT DNA-binding protein cloned by in vivo transcriptional interference

RT selection";

RL J. Bacteriol. 175:4025-4035(1993).

CC This SWISS-PROT entry is copy right. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC used by non profit institutions. There are no restrictions on its

CC modification and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).

CC send an email to license@isb-sib.ch).

CC

DR AP003596; BAB76692\_1;

DR EMBL; LL0036; AA004182\_1;

KW Hypothetical protein; Transmembrane; Signal; Complete proteome.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 436 HYPOTHETICAL PROTEIN ALR4393.

FT TRANSMEM 45 67 POTENTIAL.

SQ SEQUENCE 496 AA; 56375 MW;

454D67CF420AFB9E CRC64;

Qy 176 ADLKGSVAVSLREEIRFLPLAEV----RLAHNOQSLSLTAAIEELKTIRDLSLRDEIGQL 230

Db 177 KALQGAYEVQLKTERRAYAEAKKAIEQDRELANRQIAIEQDRELANRQOALQQRDQKI 236

Qy 231 SQLSKLTS--QIALORKE 247

Db 237 SQLDKIIQNRNLTEIA-QREE 255

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 38 IASLILIGTIGFLALGLHLVG-----FLAPQTIVTLIALFTSLAGNALYLOKTT 87

Db 8 IATIILIG- GVIATVGDITGTRYKARSLFLNRPKTVAVL---VYLITGLVLASATTI 61

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 38 IASLILIGTIGFLALGLHLVG-----FLAPQTIVTLIALFTSLAGNALYLOKTT 87

Db 8 IATIILIG- GVIATVGDITGTRYKARSLFLNRPKTVAVL---VYLITGLVLASATTI 61

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 176 ADLKGSVAVSLREEIRFLPLAEV----RLAHNOQSLSLTAAIEELKTIRDLSLRDEIGQL 230

Db 177 KALQGAYEVQLKTERRAYAEAKKAIEQDRELANRQIAIEQDRELANRQOALQQRDQKI 236

Qy 231 SQLSKLTS--QIALORKE 247

Db 237 SQLDKIIQNRNLTEIA-QREE 255

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 38 IASLILIGTIGFLALGLHLVG-----FLAPQTIVTLIALFTSLAGNALYLOKTT 87

Db 8 IATIILIG- GVIATVGDITGTRYKARSLFLNRPKTVAVL---VYLITGLVLASATTI 61

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 176 ADLKGSVAVSLREEIRFLPLAEV----RLAHNOQSLSLTAAIEELKTIRDLSLRDEIGQL 230

Db 177 KALQGAYEVQLKTERRAYAEAKKAIEQDRELANRQIAIEQDRELANRQOALQQRDQKI 236

Qy 231 SQLSKLTS--QIALORKE 247

Db 237 SQLDKIIQNRNLTEIA-QREE 255

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 38 IASLILIGTIGFLALGLHLVG-----FLAPQTIVTLIALFTSLAGNALYLOKTT 87

Db 8 IATIILIG- GVIATVGDITGTRYKARSLFLNRPKTVAVL---VYLITGLVLASATTI 61

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 176 ADLKGSVAVSLREEIRFLPLAEV----RLAHNOQSLSLTAAIEELKTIRDLSLRDEIGQL 230

Db 177 KALQGAYEVQLKTERRAYAEAKKAIEQDRELANRQIAIEQDRELANRQOALQQRDQKI 236

Qy 231 SQLSKLTS--QIALORKE 247

Db 237 SQLDKIIQNRNLTEIA-QREE 255

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 38 IASLILIGTIGFLALGLHLVG-----FLAPQTIVTLIALFTSLAGNALYLOKTT 87

Db 8 IATIILIG- GVIATVGDITGTRYKARSLFLNRPKTVAVL---VYLITGLVLASATTI 61

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 176 ADLKGSVAVSLREEIRFLPLAEV----RLAHNOQSLSLTAAIEELKTIRDLSLRDEIGQL 230

Db 177 KALQGAYEVQLKTERRAYAEAKKAIEQDRELANRQIAIEQDRELANRQOALQQRDQKI 236

Qy 231 SQLSKLTS--QIALORKE 247

Db 237 SQLDKIIQNRNLTEIA-QREE 255

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 38 IASLILIGTIGFLALGLHLVG-----FLAPQTIVTLIALFTSLAGNALYLOKTT 87

Db 8 IATIILIG- GVIATVGDITGTRYKARSLFLNRPKTVAVL---VYLITGLVLASATTI 61

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR 118

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 176 ADLKGSVAVSLREEIRFLPLAEV----RLAHNOQSLSLTAAIEELKTIRDLSLRDEIGQL 230

Db 177 KALQGAYEVQLKTERRAYAEAKKAIEQDRELANRQIAIEQDRELANRQOALQQRDQKI 236

Qy 231 SQLSKLTS--QIALORKE 247

Db 237 SQLDKIIQNRNLTEIA-QREE 255

Qy 130 DLSAV-----SQDPYSCLQGFDLSDYDQKYSTEEENRKLFSEII 175

Db 119 DKKQYETORDQAKKEKLAQODQDIAQTQAQYQORTOSRLGVVTOYQKAIAELOSYYQR-- 176

Qy 88 ANLH-----YQDLOQREVGSKKEINFLMSVLOK-----BFLHLSKEFATTSK 129

Db 62 AILFADESLRKGYFEELDQKDRKRE--QLKVAAEEQTKTQEIERNKVNQELETTR

RP	SEQUENCE OF 1-281. MEDLINE=66313706; PubMed=3018754; White A., Buckler-White A., Blumentals I., Glass R., RX	Kondo M., Sutou S.; Cloning and molecular characterization of cDNA encoding a mouse autoantigen. " RT
RA	Gorizia L.M., Hoshino Y., Floris J., Kapikian A.Z., Chanock R.M.; Conservation of amino acid sequence of VP8 and cleavage region of 84-kDa outer capsid protein among rotaviruses recovered from asymptomatic neonatal infection. Proc. Natl. Acad. Sci. U.S.A. 83:7039-7043(1986).	RT DNA Seq. 7:1-82(1997)
RA	-1- SUBCELLULAR LOCATION: Outer capsid.	RL CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR TESTIS DEVELOPMENT. PROBABLY IDENTICAL WITH THE SEROLOGICALLY DETECTABLE MALE ANTIGEN (SDM).
RT	-1- PTM: VP8 IS ONE OF TWO TRYPSIN CLEAVAGE PRODUCTS OF VP4; THE OTHER PRODUCT IS VP5.	CC CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE FOUND IN SPERMATIDS DURING SPERMATOGENESIS. NO EXPRESSION IN LEYDIG CELLS, SPERMATOGLIA, OR SPERMATOCYTES.
CC	-1- SIMILARITY: HIGH TO HUMAN GOLG1N-160.	CC CC -1- SIMILARITY: HIGH TO HUMAN GOLG1N-160.
CC	-1- CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-19 OR MET-30 IS THE INITIATOR.	CC CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-19 OR MET-30 IS THE INITIATOR.
DR	PIR; G28839; VXRW9.	CC CC
DR	DR Interpro; IPR00416; Cap_VP4.	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	PFam; PF00426; VP4; 1.	CC CC
KW	Coat protein; Glycoprotein.	CC DR
FT	CHAIN 1 776 OUTER CAPSID PROTEIN VP4.	HSSP; P18852; ISCC.
FT	CHAIN 248 776 OUTER CAPSID PROTEIN VP5.	DR
FT	CARBODY 32 32 N-LINED (GLCNAC. . .) (POTENTIAL).	MGI; MGI_96958; Colga3.
FT	CARBODY 56 56 N-LINED (GLCNAC. . .) (POTENTIAL).	DR
FT	CARBODY 97 97 N-LINED (GLCNAC. . .) (POTENTIAL).	DR
FT	CARBODY 111 111 N-LINED (GLCNAC. . .) (POTENTIAL).	DR
FT	CARBODY 114 114 N-LINED (GLCNAC. . .) (POTENTIAL).	DR
FT	CARBODY 132 132 N-LINED (GLCNAC. . .) (POTENTIAL).	Spermatogenesis; Developmental protein.
FT	CARBODY 193 193 N-LINED (GLCNAC. . .) (POTENTIAL).	KW
FT	CARBODY 278 278 N-LINED (GLCNAC. . .) (POTENTIAL).	FT DOMAIN 201 204 P0X-ALA;
FT	CARBODY 325 325 N-LINED (GLCNAC. . .) (POTENTIAL).	SEQUENCE 1325 AA; 149880 MW;
FT	CARBODY 584 584 N-LINED (GLCNAC. . .) (POTENTIAL).	3230636962C687B0 CRC64;
FT	CARBODY 607 607 N-LINED (GLCNAC. . .) (POTENTIAL).	Query Match 8.6%; Score 115; DB 1; Length 1325;
FT	CARBODY 190 190 MISSING (IN REF 2).	Best Local Similarity 21.4%; Pred. No. 4.9; Mismatches 83; Indels 42; Gaps 8;
SEQUENCE	776 AA; 87794 MW; CFCADF4DE986512 CRC64;	Matches 48; Conservative 51; Mismatches 83; Indels 42; Gaps 8;
SQ		Query Match 8.7%; Score 116; DB 1; Length 776;
	Best Local Similarity 25.8%; Pred. No. 2.2; Mismatches 31; Indels 32; Gaps 7;	Query Match 8.7%; Score 116; DB 1; Length 776;
	Matches 48; Conservative 31; Mismatches 31; Indels 32; Gaps 7;	Best Local Similarity 25.8%; Pred. No. 2.2; Mismatches 31; Indels 32; Gaps 7;
QY	94 QDLQREVGSLIKEINFMFLSLOKEPLHLSKEFATTSK DLSAYSQDFEVSCLQGFRDNYKGF 152	QY 68 VLLALFITSLAGNALYLQKTAHLHYDQLOREVGSSLKEINFMFLSLOKEPLHLSKEFATT 127
Db	492 QDLERQLGDIRE-----EFNSLSCQEIAMTQLIDALPFLDMSM----- FSGI 534	Db 610 IVLEYVALQSAKSDKEELDRGAR -RLEDIEETTSGLLEQLRQLDKVNSYEHQQTATL 668
QY	153 ESLIDEYKNTSEEMRKLFQSOEITADLGKGVASLREIRFLPLAEEYPR -- LAHNQESTL 209	Qy 128 SKDLASVSDOFYS --CLOQFRDNVKGFFESLDEYKNTSE -- EMRKLFSQELI --- 175
Db	535 KSTIDAKSAHKATVKKFKR ----- SGLATSISETLRLNSAASSVRSRSSRSNISI 588	Db 669 RQKMQVKEQPVQKVWEYRDATAKSQDNLNEKATKRLDESEMKEI -RQELIKLQSE 727
QY	210 ---TAAIELKTIDSLRDEIGOLSQLSKTLTSQALQKRESS -DLSQSQIRTLSSSPR 264	Qy 176 -----ADLKGSVASLREE -RELTPLAEEVRLAHNQESTLAI 213
Db	589 SVWTDVSEQITGSSDSVRNISTQTSASRSLRLRETTQTEGMNFNTDISAVLKTKDST 648	Db 728 KKTVEYEHSHRLQKDMSLVYHQMAELGHQSVOKE --RDMEETHLQSRLKFDEKOMIALT 784
QY	265 SASPST 270	RESULT 7
Db	649 HISPD7 654	VP4_ROTPG STANDARD; PRT; 775 AA.
QY	6	ID VP4_ROTPG STANDARD; PRT; 775 AA.
Db	G160_MOUSE	AC P23045; Created)
ID	P59371; AC 01-NOV-1997 (Rel. 20, Last sequence update)	DT 01-NOV-1991 (Rel. 20, Last annotation update)
AC	35, Created)	DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT	01-NOV-1997 (Rel. 35, Last sequence update)	DE Outer capsid protein VP4 (Hemaggutinin) [Outer layer protein VP4]
DT	15-DEC-1998 (Rel. 37, Last annotation update)	DE [Contains: Outer capsid proteins VP5 and VP8].
DT	15-DEC-1998 (Rel. 37, Last annotation update)	GN S4.
DE	Golgin-160 (Male-enhanced antigen-2) (MEA-2).	OS Porcine rotavirus (strain Gottfried).
DE	Golgin-160 (MEA-2).	OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
DE	GolgiA3 OR MEA2.	NCBI_TAXID=10317; OX NCBI_TAXID=10317;
OS	Mus musculus (Mouse).	RN [1] RN SEQUENCE FROM N.A.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	RP MEDLINE=90080150; PubMed=2152826;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	RX RX Gorziglia M., Nishikawa K., Hoshino Y., Taniguchi K.;
NCBI_TAXID=10090;	NCBI_TAXID=10090;	RT "similarity of the outer capsid protein VP4 of the Gottfried strain of porcine rotavirus to that of asymptomatic human rotavirus strains."
RN	SEQUENCE FROM N.A.	RT J. Virol. 64:414-418(1990).
RP	STRAIN=CD-1; TISSUE=Testis;	RL MEDLINE=97217683; PubMed=9063644;
RC		RX

CC -I- SUBCELLULAR LOCATION: Outer capsid.  
 CC -I- PTM: VP8 IS ONE OF TWO TRYPSIN CLEAVAGE PRODUCTS OF VP4; THE  
 CC OTHER PRODUCT IS VP5.

CC -I- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).

DR EMBL: M33516; AAA47095; 1; - .

DR PTR: A33563; VPXRGP.

DR Pfam: PF00426; VP4; 1.

KW Coat protein; Glycoprotein.

FT CHAIN 1: 775 OUTER CAPSID PROTEIN VP4.  
 FT CHAIN 1: 240 OUTER CAPSID PROTEIN VP5.  
 FT CHAIN 1: 775 OUTER CAPSID PROTEIN VP5.

FT CARBOHYD 32: 32 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 56: 56 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 97: 97 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 132: 132 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 324: 324 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 583: 583 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 775 AA: 87098 MW: C8C736A85CD03275 CRC64; .

Query Match 8.6%; Score 114; DB: 1; Length 775;  
 Best Local Similarity 23.3%; Pred. No. 3;  
 Matches 51; Conservative 41; Mismatches 91; Indels 36; Gaps 8;

Qy 61 IAFQITIVLLALFTSLAGNALYQLQTAN-LHLYQDLOREVGSLKREINFMLSVLQKEFLH 119

Db 463 IAGRFSLLL-----VPSNDYQPTIMNSVTVRQDLERQGDLRE-----EFTN 506

Qy 120 LSKEFATSK-DLSAVSQDFYSCLOQFRDNVKGFESLLDKNSSTEEMRKLFSQE-----I 174

Db 507 LSQEIAMTQLIDALLPLDMFSM-----FSGIKRSTIDYAKSATVNKKFKSGLAYS 559

Qy 175 IADLKGSVASLREEIRFLPLAEVRRLAHQESTAAVFLKTIRDSLRDEIGQLSQLI 234

Db 560 ISELTSLSPSASSVSRSSRSRNISSS----VWTDVSEQIADASNSVRSISTQTSATS 615

Qy 235 KTLTSQIQLAKRES--SDLSQSQRETLLSPRKASASP 270

Db 616 KRLRLETTTQTEGMNFDDISAAVLKTPLDKSTHISPD 654

RESULT 8  
 LAMO\_DROME  
 ID LAMO\_DROME STANDARD: PRT; 622 AA.  
 AC P08948;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-OCT-1996 (Rel. 33, Last sequence update)  
 DE LamIn Dm0.  
 GN LAM  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydriodea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=727;

RN SEQUENCE FROM N.A.  
 RP MEDLINE-91065636; PubMed=2123469;  
 RX Oshan M.; Paz M.; Landesman Y.; Fainsod A.; Gruenbaum Y.;  
 RT "Molecular analysis of the Drosophila nuclear lamin gene.";  
 RL Genomics 8:217-224 (1990).  
 RN SEQUENCE FROM N.A.

RX MEDLINE=88153889; PubMed=3126192;  
 RA Gruenbaum Y., Landesman Y., Drees B., Bare J.W., Saumweber H.,  
 RA Paddy M.R., Sedat J.W., Smith D.E., Benton B.M., Fisher P.A.;  
 RT "Drosophila nuclear lamin precursor Dm0 is translated from either of  
 two developmentally regulated mRNA species apparently encoded by a  
 single gene.";  
 RT J. Cell Biol. 106:585-596(1988).  
 RN [3]  
 RP REVISIONS TO 24-39 AND 254-263.  
 RA Stuurman N., Maus N., Fisher P.A.;  
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.  
 CC -I- FUNCTION: LAMINS ARE COMPONENTS OF THE NUCLEAR LAMINA, A FIBROUS  
 LAYER ON THE NUCLEOPLASMIC SIDE OF THE INNER NUCLEAR MEMBRANE,  
 WHICH IS THOUGHT TO PROVIDE A FRAMEWORK FOR THE NUCLEAR ENVELOPE,  
 AND MAY ALSO INTERACT WITH CHROMATIN.  
 CC -I- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -I- PTM: THREE FORMS OF LAMIN HAVE BEEN IDENTIFIED IN D.MELANOGASTER,  
 CC Dm1 IS RAPIDLY PROCESSED TO LAMIN DM1 IN THE CYTOPLASM,  
 CC PHOSPHORYLATED, FORMING LAMIN DM2.  
 CC -I- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.  
 CC THIS IS A B TYPE LAMIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).

DR EMBL: X16275; CAA34351; 1; - .

DR PIR: X02728; CAA34359; 1; - .

DR PIR: A29965; A29965;

DR PIR: A37103; A37103;

DR Flybase: FBgn0002525; Lam.

DR InterPro: IPR001664; IF.

DR InterPro: IPR001322; IF.

PFam: PF00038; filament, 1.

DR Pfam: PF00932; filament, 1.

DR PROSITE; PS00226; IF; FALSE NEG.

KW Intermediate filament; Colled coil; Nuclear protein; Lipoprotein;

KW Phosphorylation; Head.

FT DOMAIN 1: 56 DOMAIN 1: 56

FT DOMAIN 1: 57 DOMAIN 1: 408

FT DOMAIN 1: 409 DOMAIN 1: 408

FT DOMAIN 1: 55 DOMAIN 1: 622

FT DOMAIN 1: 91 DOMAIN 1: 622

FT DOMAIN 1: 103 DOMAIN 1: 622

FT DOMAIN 1: 104 DOMAIN 1: 622

FT DOMAIN 1: 241 DOMAIN 1: 622

FT DOMAIN 1: 242 DOMAIN 1: 622

FT DOMAIN 1: 265 DOMAIN 1: 622

FT DOMAIN 1: 266 DOMAIN 1: 622

FT SITE 289 SITE 289

FT SITE 353 SITE 353

FT DOMAIN 446 DOMAIN 446

FT LIPID 619 LIPID 451

FT CONFLICT 24 CONFLICT 24

FT CONFLICT 24 CONFLICT 39

FT CONFLICT 135 CONFLICT 39

FT CONFLICT 255 CONFLICT 255

FT CONFLICT 270 CONFLICT 270

SQ SEQUENCE 622 AA: 71249 MW: DCBF0187CCF2E2A8 CRC64;

Query Match 8.4%; Score 112.5; DB: 1; Length 622;  
 Best Local Similarity 24.3%; Pred. No. 2, 9; Mismatches 39; Indels 33; Gaps 8;  
 Matches 50; Conservative 39; Gaps 8;

Qy 75 TSLAGNALLYQQTANLHYQDLOREVGSLKEINFMSVLYQKEFHLHSKEFATPSKDSLAV 134

Db 151 TTARGNVRMYESEN-ELNKNKYQANDRKKLNEDEALKELERLQFEETRKNEEQE 209

Qy 135 S---QDFYSCLOQFRDNYKGFFESLLDEYKNSTEMRKL-----FSQLIADIKGGSVA 183

Qy : ---QDFYSCLOQFRDNYKGFFESLLDEYKNSTEMRKL-----FSQLIADIKGGSVA 183

Db 210 TLSRDLENTIQSLREELSFKDOIHSQEINERSRIKOTYESEIDGRSLSEYDAKLKQSLQ 269  
 DR PIR: C24700; C24700.  
 DR HSSP; P02671; 1FZB.  
 DR InterPro; IPR00074; Apolipoprotein.  
 DR PF01442; Apolipoprotein; 1.  
 DR Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal.  
 KW 1 20  
 FT SIGNAL  
 FT CHAIN 21 391  
 FT DOMAIN 33 330  
 FT REPEAT 33 54 1.  
 FT REPEAT 60 81 2.  
 FT REPEAT 82 103 3.  
 FT REPEAT 115 136 4.  
 FT REPEAT 137 158 5.  
 FT REPEAT 159 180 6.  
 FT REPEAT 181 202 7.  
 FT REPEAT 203 224 8.  
 FT REPEAT 225 246 9.  
 FT REPEAT 247 268 10.  
 FT REPEAT 269 286 11.  
 FT REPEAT 287 308 12.  
 FT REPEAT 309 330 13.  
 FT DOMAIN 374 385 14.  
 FT VARIANT 253 Q->H.  
 SQ SEQUENCE 391 AA; 4456 MW; 24095004A809201D CRC64;  
 [1]  
 RN STANDARD; PRT; 391 AA.  
 ID APA4\_RAT  
 AC P02651;  
 AC 21-JUL-1986 (Rel. 01, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 RT Apolipoprotein A-IV precursor (Apo-AIV).  
 OS Rattus norvegicus (Rat).  
 OC Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Muridae; Murinae; Rattus .  
 OX NCBI\_Taxid=10116;  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE-86196059; PubMed=3009456;  
 RX Boguski M.S., Birkenmeier E.H., Elshourbagy N.A., Taylor J.M.,  
 RA Gordon J.I.;  
 RT "Evolution of the apolipoproteins. Structure of the rat apo-A-IV gene and its relationship to the human genes for apo-A-I, C-III, and E.";  
 RL J. Biol. Chem. 261:6398-6407(1986).  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE-84298874; PubMed=6591177;  
 RX Boguski M.S., Elshourbagy N.A., Taylor J.M., Gordon J.I.,  
 RA "Rat apolipoprotein A-IV contains 13 tandem repetitions of a 22-amino acid segment with amphipathic helical potential.";  
 RT and its relationship to the human genes for apo-A-I, C-III, and E.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5021-5025(1984).  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE-87008540; PubMed=3020028;  
 RX Haddad I.A., Ordovas J.M., Fitzpatrick T., Karathanasis S.K.;  
 RA RT "Linkage, evolution, and expression of the rat apolipoprotein A-I, C-III, and A-IV genes";  
 RL J. Biol. Chem. 261:13268-13277(1986).  
 CC FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND CATABOLISM, REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN LIPASE BY APOC-II, POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR COMPONENT OF HDL AND CHYLOMICRONS.  
 CC SUBCELLULAR LOCATION: Extracellular.  
 CC TISSUE SPECIFICITY: SECRETED IN PLASMA.  
 CC DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH 1-  
 CC 22 MER) IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-  
 CC MEMBERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-  
 CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY  
 CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN/CHOLESTEROL  
 ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.  
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOB FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; J02588; AAA07441; -.  
 DR EMBL; M13508; AAA074481; -.  
 DR PIR; A03095; LPRTAA.  
 DR PIR; A25214; A25214.  
 DR



[3]	IRN	SEQUENCE OF 1-8 FROM N.A.
RP	Bai, Y.	Symington, L.S.
RA	Submitted (May-1996) to the EMBL/GenBank/DDBJ databases.	
CC	-1- FUNCTION REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI COMPLEX.	
CC	-1- SUBCELLULAR LOCATION: CITOPLASMIC; ASSOCIATED WITH INTRACELLULAR MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATING BETWEEN THE ER AND THE GOLGI COMPLEX.	
CC	-1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.	
CC	-1- SIMILARITY: BELONGS TO THE VIP/US01/YB047C FAMILY.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce">http://www.isb-sib.ch/announce</a> or send an email to license@isb-sib.ch).	
CC	EMBL: X54378; CAA39253; 1; -. EMBL: L03188; AAB00143; 1; -. EMBL: U53668; AAB66659; 1; -. PIR: A38455; A38455. HSSP: P80240; 1DIP. SGD: S0002216; US01.	
CC	InterPro: IPR002017; Spectrin. IntProt: PRO002017; Spectrin. Transport: Protein transport; Golgi stack; Cytoskeleton; Coiled coil. TRN: 724 KW: GLOBULAR HEAD. TRN: 1 KW: COILED COIL (POTENTIAL). TRN: 725 KW: CHARGED (HYPER-HYDROPHILIC). TRN: 465 KW: DISPENSABLE FOR THE PROTEIN FUNCTION. TRN: 991 KW: ASP/GLU-RICH (ACIDIC). TRN: 1172 KW: GLU-RICH (ACIDIC). TRN: 847 KW: G > E (IN REF. 2). TRN: 924 KW: E -> K (IN REF. 2). TRN: 1253 KW: V >> I (IN REF. 2). TRN: 1319 KW: I >> V (IN REF. 2). TRN: 1461 KW: S (IN REF. 2). TRN: 1581 KW: G >> S (IN REF. 2). TRN: 1600 KW: I >> V (IN REF. 2). TRN: 1661 KW: R >> S (IN REF. 2). TRN: 1772 KW: D >> DEEDDEE (IN REF. 2). SEQUENCE 1790 AA; 206424 MW;	
CC	62C2B2169FD4818 CRC64; SQ	
CC	Query Match Score 8.4%; Score 111.5; DB 1; Length 1790; Best Local Similarity 22.3%; Pred. No. 12; Mismatches 81; Indels 65; Gaps Matches 54; Conservative 42; Mismatches 81; Indels 65; Gaps	
Qy	74 ITSLAGNALYQKTAHNLHYQDQLQREVGLS-----KEINFMLSVLQKEFLHLSKE-- 1.23	
Db	964 LKSLANN-----YKDMQAENESLIKAVEEESKNNSIQLSMLQNKTDSMSQEKE 1011	
Qy	124 -----PATTSKDLSAVSQDFYSCLQGRDNKYKGFESL-----D 157	
Db	1012 NFQIERSIEKNEOLKKTISDLEQTEELISKSDSSRDEVESQISLSEKE 1071	
Qy	158 EYKRNSTEEPMRQLFQSQTIAADKGSVASLREIRFLPLAEE-VRLRAHNQESL----- 209	
Db	1072 ENVKISLTKT-REELEAEAL AAYKLNKNELETKLTSEKALKVENEELKEEIQIOL 1129	
Qy	210 ----TAAIEELKTIR--DSLDRDEIGQLSQLSKTLTQIALORKESSDLCSQIRETNSSP 262	
Db	1130 EKEATEKQQLNSLRANLESLEKEHDLAQQLKYEQQANKERQYNEEISQINDELST 1189	
Qy	263 RK 264	
Db	1190 QQ 1191	

FT DOMAIN 215 230 LINER 12.  
 FT DOMAIN 231 252 COTI 2A.  
 FT DOMAIN 233 256 LINER 2B.  
 FT DOMAIN 257 377 COTI 2B.  
 SQ SEQUENCE 432 AA; 4980 MW; E6C3B3454C3F1250 CRC64;

Query Match 8.3%; Score 111; DB 1; Length 432;  
 Best local Similarity 24.4%; Pseqd. No. 2.;  
 Matches 60; Conservative 34; Mismatches 86; Indels 66; Gaps 11;

Qy 82 LYLQKTANLHYQ-----DLOREYGLS-KELNFMSVLQKEFLHLSKEPA-TT 127  
 Db 161 LRLEAENLAAAYRQEADATLARDLERKISLEEFRLRKHEEVRELQEQLARQVY 220  
 Qy 128 SKDLAYSQDFESCLQGRFDNYKGFS-----  
 Db 221 HVELDVAKPDLTAALKTRTQEAMASNMHEAEWYRSKFADLTDAARNAELRQAH 280  
 Qy 162 STEEMRKFLSQEIIAD--LKGSVASLRLERFLPLAAEVRLAHNOESLTAATEELKT 218  
 281 EANDYRQLL-QSLTCDELSLRGTNEELRQMR-EQEERTVREASYQDALARLEEGJS 337  
 Qy 219 IRDSLRL--DBIGQSLQSKPLTSQLQALQRK-----ESSDILCSQRETLS 261  
 Db 338 LKDEMARHLQDYQDLINVKLAUDIEIATYKLLGEENRITIPQTSNL--QIRET-SL 394  
 Qy 262 PRKSAS 267  
 Db 395 DTKSVS 400

RESULT 14  
 MYS3\_SCPO STANDARD: PRT; 2104 AA.  
 ID MYS3\_SCPO STANDARD: PRT; 2104 AA.  
 AC O14157; O42730;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin type II heavy chain 2.  
 GN MY03 OR MYP2 OR SPAC4AB\_05C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycteales; Schizosaccharomyctaceae;  
 OC Schizosaccharomyces.  
 NCBI TAXID=4896;  
 RN SEQUENCE FROM N.A.; MEDLINE=98119390; PubMed=9459302;  
 RX Motegi F., Nakao K., Kitayama C., Yamamoto M., Mabuchi I.;  
 "Identification of Myo1, a second type-II myosin heavy chain in the  
 fission yeast Schizosaccharomyces pombe.", FEBS Lett. 420:161-166(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.; MEDLINE=98052346; PubMed=9398685;  
 RX Beranilla M., Poisburg S.L., Pollard T.D.;  
 "Identification of a second myosin-II in Schizosaccharomyces pombe:", Myp2p is conditionally required for cytokinesis.", Mol. Biol. Cell 8:2693-2705(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V., Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.  
 RL - FUNCTION: STABILIZES THE F ACTIN CABLES FORMING THE F-ACTIN RING  
 CC THAT SURROUNDS THE NUCLEUS DURING INTERPHASE. MAY WORK IN CONJUNCTION WITH MYO2.  
 CC SUBUNIT: BINDS TO CDC4 AND RICL.  
 CC SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; AB007633; BAA4579.1; -.  
 DR EMBL; AF023788; CAB1175.1; -.  
 DR EMBL; Z98562; CABI1175.1; -.  
 DR HSSP; P08199; IMND.  
 DR InterPro; IPR004009; Myosin\_N  
 DR InterPro; IPR001609; myosin\_head.  
 DR Pfam; PF00663; myosin\_head; 1.  
 DR PRINTS; PRO0193; MYOSIN\_N; 1.  
 DR PRODOM; PDO0355; myosin\_head; 1.  
 DR SMART; SM0242; MYC; 1.  
 KW Myosin; Coiled coil; ATP-binding; Actin-binding; Alkylation.  
 FT DOMAIN 1 828 MYOSIN HEAD-LIKE.  
 FT DOMAIN 2104 COILED COIL (POTENTIAL), ACTIN-BINDING (POTENTIAL).  
 FT DOMAIN 646 660 ACTIN-BINDING (POTENTIAL).  
 FT NP\_BIND 182 189 NP (POTENTIAL).  
 FT MOD\_RES 686 686 ALKYLYATION (BY SIMILARITY).  
 FT CONFLICT 1193 1193 D->G (IN REF. 2).  
 FT CONFLICT 1304 1304 E->G (IN REF. 2).  
 FT CONFLICT 1344 1344 E->K (IN REF. 2).  
 SQ CONFLICT 1420 1420 G->D (IN REF. 2).  
 SQ SEQUENCE 2104 AA; 242570 MW; 3A05485940280258 CRC64;

Query Match 8.3%; Score 110.5; DB 1; Length 2104;  
 Best Local Similarity 21.8%; Pred. No. 16;  
 Matches 41; Conservative 48; Mismatches 72; Indels 27; Gaps 6;  
 Qy 93 YQDLQREVGSKIEINFMLSVLOKEFFHLSKEPATTSQDLQYDYSCLQFRDNYKGF 152  
 Db 918 YFFLYAKRQSIVTE--DLANKQTEISYSLSDLSSTLEKKLSSIKKDQTISSKYKELEKDY 974  
 Qy 153 ESLLDYKNSNDEMRKLQSOETIADLQGSVASLREIFTRPLAEEYVRLAHNQESLTAA 212  
 Db 975 LNIMADYQHSSQHSLN--ERAINEKNLNIRELNELNLMRLDD----ELLLKQSYDTK 1026  
 Qy 213 TEELKTYTRDSLDEI---GOL-----SOLSKTLTSQIALQKREKSSDLCSQIRE 257  
 Db 1027 YQELRRENASTKDCQRTYESOLASLVSKEYSETESEINKK-EAELVTFQKITEYRDQLHK 1085  
 Qy 258 TLSSPKS 265  
 Db 1086 AFQNPEKT 1093

RESULT 15  
 VP4\_ROTHT STANDARD; PRT; 774 AA.  
 ID VP4\_ROTHT STANDARD; PRT; 774 AA.  
 AC F11200;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-Nov-1980 (Rel. 16, Last sequence update)  
 DT 16-Oct-2001 (Rel. 40, Last annotation update)  
 DE Outer capsid protein VP4 (Hemagglutinin) (Outer layer protein VP4)  
 DE [Contains: Outer capsid proteins VP5 and VP8].  
 GN S4.  
 OS Human rotavirus (serotype 4 / strain St. Thomas 3).  
 OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.; SEQUENCE FROM N.A.;  
 RX MEDLINE=88275070; PubMed=2839714;  
 RA Gorisigilia M., Green K., Nishikawa K., Taniguchi K., Jones R., Kapikian A.Z., Chanock R.M.;  
 RA "Sequence of the fourth gene of human rotaviruses recovered from asymptomatic or symptomatic infections.";  
 RT This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation.  
 RL J. Virol. 62:2978-2984(1988).  
 RN [2]

RP SEQUENCE OF 1-280.  
 RX MEDLINE-86311706; PubMed=3018754;  
 RA Gorillaia M., Hostino Y., Buckler-White A., Blumentals I., Glass R.,  
 RA Flores J., Kapikian A.Z., Chanock R.M.;  
 RT "Conservation of amino acid sequence of VP8 and cleavage region of  
 RT 84-kDa outer capsid protein among rotaviruses recovered from  
 asymptomatic neonatal infection.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 83:7039-7043 (1986).  
 RL -1- SUBCELLULAR LOCATION: Outer capsid.  
 CC -1- PTM: VP8 IS ONE OF TWO TRYPSIN CLEAVAGE PRODUCTS OF VP4; THE  
 CC OTHER PRODUCT IS VP5.  
 CC -1- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.  
 DR PIR; H28839; VPXRNT.  
 DR InterPro; IPR00416; cap-VP4.  
 DR Pfam; PF00426; VP4; 1.  
 KW Coat protein; Glycoprotein.  
 FT CHAIN 1 774 OUTER CAPSID PROTEIN VP4.  
 FT CHAIN 1 240 OUTER CAPSID PROTEIN VP8.  
 FT CHAIN 247 774 OUTER CAPSID PROTEIN VP5.  
 CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 583 583 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 774 AA: 87455 MW: D397E5C6125A8FA8 CRC64;

Query Match 8.3%; Score 110; DB 1; Length 774;  
 Best Local Similarity 25.3%; Pred. No. 5.4;  
 Matches 47; Conservative 31; Mismatches 76; Indels 32; Gaps 7;  
 Mismatches 47;

QY 94 QDLQREVGSLKEINFMLSVLQKFELHSKEFATTSK-DLSAVSQDFYSCLQGERFDNYKGF 152  
 Db 491 QDLERQGDLRE-----EFNSLSQLAITOOLIDALLPLDMFSM-----FSGI 533  
 QY 153 ESLLDYEKNTSTEEMRKLFQSEIADIUKGSVASLREPIRFLTPLAEVRR--LAHNOESI 209  
 Db 534 KSTIDAQAKSMATVKMKPFKR-----SGLATISSELTRLSNAASSVSRESSRSIRNISII 587  
 QY 210 --TAAEEELKTRDSURDETQLSQSKTTSQTAQRESS--DLCSCRETISSSPRK 264  
 588 SEPTIDVSEQLAGSSSDSVRNISTQTASRRRLREITTOEGMNFIDSAAVLKTKIDRST 647  
 QY 265 SASPST 270  
 Db 648 HIRPDT 653

Search completed: August 13, 2002, 09:30:44  
 Job time: 380 sec

Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 13, 2002, 09:30:16 ; Search time 48.57 Seconds

(without alignments)

972.362 Million cell updates/sec

Title:	US-09-673-763-14
Perfect score:	1333
Sequence:	1 MTTPLIVIIPPSSPAPSYSA.....QIRETLSSRKSASPSTKS 273
Scoring table:	BLOSUM62
	Gapop 10.0 , Gapext 0.5
Searched:	5622222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters:	5622222
Minimum DB seq length: 0	
Maximum DB seq length: 2000000000	
Post-processing: Minimum Match 0%	
Maximum Match 100%	
Listing first 45 summaries	
Data base :	SPTREMBL_19: 1: sp_archaea: 2: sp_bacteria: 3: sp_fungi: 4: sp_human: 5: sp_invertebrate: 6: sp_mammal: 7: sp_minc: 8: sp_organelle: 9: sp_phage: 10: sp_plant: 11: sp_rodont: 12: sp_virus: 13: sp_vertebrate: 14: sp_unclassified: 15: sp_virus: 16: sp_bacteriopl: 17: sp_archaea: 

Result No.	Score	Query Match	Length	DB ID	Description
1	1333	100.0	273	2	069196 chlamydia t
2	1328	99.6	273	2	069196 chlamydia t
3	1318	98.9	273	2	Q9AM5 chlamydia t
4	1315	98.6	273	2	Q9AMB2 chlamydia t
5	1311	98.3	273	2	Q9AMA4 chlamydia t
6	1310	98.3	273	16	Q84121 chlamydia t
7	1308	98.1	273	2	Q9AMA6 chlamydia t
8	1307	98.0	273	2	Q9AMA9 chlamydia t
9	1306	98.0	273	2	Q9AMA9 chlamydia t
10	1306	98.0	273	2	Q9F7K9 chlamydia t
11	1305	97.9	273	2	Q9AMA7 chlamydia t
12	1305	97.9	273	2	Q99Q56 chlamydia t
13	1302	97.7	273	2	Q9AMA8 chlamydia t
14	1293	97.0	273	2	Q9REFX chlamydia t
15	1292	96.9	273	2	Q9AMB0 chlamydia t
16	1209.5	90.7	256	2	Q9F7L1 chlamydia t

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100.0	273	2	069196 chlamydia t	069196 chlamydia t
2	99.6	273	2	069196 chlamydia t	069196 chlamydia t
3	98.9	273	2	Q9AM5 chlamydia t	Q9AM5 chlamydia t
4	98.6	273	2	Q9AMB2 chlamydia t	Q9AMB2 chlamydia t
5	98.3	273	2	Q9AMA4 chlamydia t	Q9AMA4 chlamydia t
6	98.3	273	16	Q84121 chlamydia t	Q84121 chlamydia t
7	98.1	273	2	Q9AMA6 chlamydia t	Q9AMA6 chlamydia t
8	98.0	273	2	Q9AMA9 chlamydia t	Q9AMA9 chlamydia t
9	98.0	273	2	Q9F7K9 chlamydia t	Q9F7K9 chlamydia t
10	98.0	273	2	Q9AMA7 chlamydia t	Q9AMA7 chlamydia t
11	97.9	273	2	Q99Q56 chlamydia t	Q99Q56 chlamydia t
12	97.9	273	2	Q9AMA8 chlamydia t	Q9AMA8 chlamydia t
13	97.7	273	2	Q9REFX chlamydia t	Q9REFX chlamydia t
14	97.0	273	2	Q9AMB0 chlamydia t	Q9AMB0 chlamydia t
15	96.9	273	2	Q9AMBI chlamydia t	Q9AMBI chlamydia t
16	90.7	256	2	Q9F7L1 chlamydia t	Q9F7L1 chlamydia t

Result No.	Score	Query Match	Length	DB ID	Description
1	96.2	72.2	197	2	Q9F7T6 chlamydia t
2	756.5	56.8	158	2	Q9F7L0 chlamydia t
3	714	53.6	276	16	Q9PRR8 chlamydia t
4	612	45.9	174	2	Q9F712 chlamydia t
5	582	43.7	121	2	Q9AM93 chlamydia t
6	564	42.3	174	2	Q9F7K8 chlamydia t
7	503	37.7	109	2	Q9F7T7 chlamydia t
8	162.5	12.2	355	2	Q46210 chlamydophi
9	154.5	11.6	390	16	Q9BZ88 chlamydia t
10	128.5	9.6	472	24	Q07952 carassius p
11	126	9.5	1179	17	Q59462 pyrococcus
12	124	9.3	263	13	Q98t6 anguilla ja
13	124	9.3	347	13	Q02428 saccharomy
14	123.5	9.3	471	13	P87360 brachydienio
15	122.5	9.2	355	4	Q14789 homo sapien
16	122	9.2	775	12	Q9f7E8 chlamydia t
17	122	9.2	1931	5	Q86200 human rotav
18	124	9.0	511	13	Q9ncF9 drosophila
19	124	9.0	900zF7 rana catesbe		
20	124	9.0	9by13 homo sapien		
21	120.5	9.0	120.5	5	Q21025 caenorhabdi
22	120	9.0	644	16	Q9Km20 vibrio chol
23	120	9.0	1931	5	Q9VKH9 drosophila
24	119.5	9.0	406	2	Q9L5T4 salmonella
25	119.5	9.0	1743	5	Q96063 dugesia jap
26	119.5	8.9	897	13	Q13098 xenopus lae
27	117.5	8.8	1475	5	Q90314 human rotav
28	117.5	8.8	1475	5	Q76447 caenorhabdi
29	117	8.8	1101	4	Q96SB9 homo sapien

## ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	1	069196	1	ID 069196	PRELIMINARY;
				AC 069196;	PRM; 273 AA.
				DT 01-AUG-1998	(TREMBLrel. 07, Created)
				DT 01-AUG-1998	(TREMBLrel. 07, Last sequence update)
				DT 01-DEC-2001	(TREMBLrel. 19, Last annotation update)
				DE INCLUSION PROTEIN.	
				GN INCA.	
				OS Chlamydia trachomatis.	
				OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.	
				NCBI_TaxID=813;	
				RN [1]	
				RP SEQUENCE FROM N.A.	
				RC STRAIN=LGV-434;	
				RX MEDLINE#99043938;	
				RA Bannantine J.P.; Stamm W.E.; Sichland R.J.; Rockey D.D.;	
				RT "Chlamydia trachomatis Inca is localized to the inclusion membrane and is recognized by antisera from infected humans and primates.";	
				RL Infect. Immun. 66:6017-6021(1998).	
				DR AF067958; AAC82641; -	
				SQ SEQUENCE 273 AA; 30272 MW; E64BB46A5D6824 EC CRC64;	

Query Match	Best Local Similarity	Length	DB 2;
			Pred. No. 1.6e-81;
			Mismatches 0;
			Gaps 0;
1	MTTP7LIVTPSPPPAPSYSSANRVPQPSLMDKIKITATASLILGTGFLLALGHVGFL	60	
2	1MTTP7LIVTPSPPPAPSYSSANRVPQPSLMDKIKITATASLILGTGFLLALGHVGFL	60	
3	IAPOITIVVLLALPFTSLAGNALYLQKTAHLYDQREVGSLKEINFMVLSQLKEFHL	120	
4	IAPOITIVVLLALPFTSLAGNALYLQKTAHLYDQREVGSLKEINFMVLSQLKEFHL	120	
5	Q9AM9 chlamydia t	60	
6	Q9f7k9 chlamydia t	60	
7	Q9AMA4 chlamydia t	60	
8	Q9AMA6 chlamydia t	60	
9	Q9AMA9 chlamydia t	60	
10	Q9AMA7 chlamydia t	60	
11	Q9AMA8 chlamydia t	60	
12	Q9AMA9 chlamydia t	60	
13	Q9AMA10 chlamydia t	60	
14	Q9AMA11 chlamydia t	60	
15	Q9AMA12 chlamydia t	60	
16	Q9AMA13 chlamydia t	60	
17	Q9AMA14 chlamydia t	60	
18	Q9AMA15 chlamydia t	60	
19	Q9AMA16 chlamydia t	60	
20	Q9AMA17 chlamydia t	60	
21	Q9AMA18 chlamydia t	60	
22	Q9AMA19 chlamydia t	60	
23	Q9AMA20 chlamydia t	60	
24	Q9AMA21 chlamydia t	60	
25	Q9AMA22 chlamydia t	60	
26	Q9AMA23 chlamydia t	60	
27	Q9AMA24 chlamydia t	60	
28	Q9AMA25 chlamydia t	60	
29	Q9AMA26 chlamydia t	60	
30	Q9AMA27 chlamydia t	60	
31	Q9AMA28 chlamydia t	60	
32	Q9AMA29 chlamydia t	60	
33	Q9AMA30 chlamydia t	60	
34	Q9AMA31 chlamydia t	60	
35	Q9AMA32 chlamydia t	60	
36	Q9AMA33 chlamydia t	60	
37	Q9AMA34 chlamydia t	60	
38	Q9AMA35 chlamydia t	60	
39	Q9AMA36 chlamydia t	60	
40	Q9AMA37 chlamydia t	60	
41	Q9AMA38 chlamydia t	60	
42	Q9AMA39 chlamydia t	60	
43	Q9AMA40 chlamydia t	60	
44	Q9AMA41 chlamydia t	60	
45	Q9AMA42 chlamydia t	60	

Query Match	Best Local Similarity	Length	DB 2;
			Length 273;
			Matches 273;
			Conservative 0;
			Mismatches 0;
			Gaps 0;
Qy	1 MTTP7LIVTPSPPPAPSYSSANRVPQPSLMDKIKITATASLILGTGFLLALGHVGFL	60	
Db	1 MTTP7LIVTPSPPPAPSYSSANRVPQPSLMDKIKITATASLILGTGFLLALGHVGFL	60	
Qy	61 IAPOITIVVLLALPFTSLAGNALYLQKTAHLYDQREVGSLKEINFMVLSQLKEFHL	120	
Db	61 IAPOITIVVLLALPFTSLAGNALYLQKTAHLYDQREVGSLKEINFMVLSQLKEFHL	120	
Qy	121 SKERATSKDLDAVSDPFYSCLOGFRDNKGFFESLDEYKNSSTEEMRKLFSEIIADLKG	180	
Db	121 SKERATSKDLDAVSDPFYSCLOGFRDNKGFFESLDEYKNSSTEEMRKLFSEIIADLKG	180	



Qy	121	SKEFATSKDLSAVSQDFYSCLOGFRDNYKGFESELDEYKNSDEEMRKLFSEBLIADIKG 180		Query Match Score 98.3%; Best Local Similarity 98.2%; Pred. No. 4; Gaps 0;
Db	121	SKEFATSKDLSAVSQDFYSCLOGFRDNYKGFESELDEYKNSDEEMRKLFSEBLIADIKS 180		Matches 268; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy	181	SVASLREIRFLPLAEEYRRLAHNOESTAAEELKLIRDSELDEIGOLSOLSKTISQ 240		Qy 1 MTTPTLIVIPSPPPAPSYANRPQPSLMDKIKKTAIAASLILIGTIGFALLGHLVGFL 60
Db	181	SVASLREIRFLPLAEEYRRLAHNRESLTAAEELKLIRDSELDEIGOLSOLSKTISQ 240		Db 1 MTTPTLIVIPSPPPAPSYANRPQPSLMDKIKKTAIAASLILIGTIGFALLGHLVGFL 60
Qy	241	IALQRKESSDLCSQIRETLSSPRKSASSSTKSS 273		Qy 61 IAPQTIVVLLAFTSLAGNLYLQTANLHYLQDOREVGSKETINFMLSVLQKEFLHL 120
Db	241	IALQRKESSDLCSQIRETLSSPRKSASSSTKSS 273		Db 61 IAPQTIVVLLAFTSLAGNLYLQTANLHYLQDOREVGSKETINFMLSVLQKEFLHL 120
RESULT	5			Qy 121 SKEFATSKDLSAVSQDFYSCLOGFRDNYKGFESELDEYKNSDEEMRKLFSEBLIADIKG 180
RP	084121	PRELIMINARY; PRT; 273 AA.		Db 121 SKEFATSKDLSAVSQDFYSCLOGFRDNYKGFESELDEYKNSDEEMRKLFSEBLIADIKG 180
RC	084121;			Qy 181 SVASLREIRFLPLAEEYRRLAHNOESTAAEELKLIRDSELDEIGOLSOLSKTISQ 240
DT	01-NOV-1998	(TREMBLrel. 08, Created)		Db 181 SVASLREIRFLPLAEEYRRLAHNRESLTAAEELKLIRDSELDEIGOLSOLSKTISQ 240
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)		Qy 241 IALQRKESSDLCSQIRETLSSPRKSASSSTKSS 273
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		Db 241 IALQRKESSDLCSQIRETLSSPRKSASSSTKSS 273
N		INCLUSION MEMBRANE PROTEIN A.		
OS		Chlamydia trachomatis.		
OC		Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.		
OX		NCBI_TaxID:813;		
RN		SEQUENCE FROM N.A.		
RC		STRAIN=D/DW 3/CX.		RN [1] RP SEQUENCE FROM N.A.
RX		MEDLINE=99000809; PubMed=9784136;		RC ID Q9AM46 PRELIMINARY; PRT; 273 AA.
RA		Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,		AC Q9AM46;
RA		Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,		DT 01-JUN-2001 (TREMBLrel. 17, Created)
RA		"Genome sequence of an obligate intracellular pathogen of humans:		DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
RT		Chlamydia trachomatis."		DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
RT		Science 282:754-759(1998).		DE INCLUSION MEMBRANE PROTEIN A.
RL		[2]		GN INCA.
RN		SEQUENCE FROM N.A.		OS Chlamydia trachomatis.
RP		SEQUENCE FROM N.A.		OC NCBITaxonID=813;
RA		Virayosin W., Rockey D.D., Suchland R.J., Stamm W.E.,		RN [1] RN [2] RP SEQUENCE FROM N.A.
RT		"Diverse Mutations in Inca Amplified from Clinical Chlamydia trachomatis Isolates that Occupy Nonfusogenic Inclusions.";		RC STRAIN=D/40-L;
RL		Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.		RX MEDLINE=21295121; PubMed=11402010;
RN		[3]		RA Pannekoek Y., van Der Ende A., Eijk P.P., van Marle J., de Witte M.A.,
RP		SEQUENCE FROM N.A.		RA Ossewaarde J.M., van Den Brule A.J.C., Morre S.A., Dankert J.;
RC		STRAIN=M79301;		RT "Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the Inca 147T mutation.";
RA		Virayosin W., Rockey D.D., Suchland R.J., Stamm W.E.,		RT Infect. Immun. 69:4654-4656 (2001).
RT		"Diverse Mutations in Inca Amplified from Clinical Chlamydia trachomatis Isolates that Occupy Nonfusogenic Inclusions.";		DR EMBL; AF327008; AAC61105.1; -.
RL		[4]		SQ SEQUENCE 273 AA; 30265 MW; 2267957C555E572B CRC64;
N		SEQUENCE FROM N.A.		
RC		STRAIN-SA-1, APACHE-2, IOL-238, UW-36, 2B, 10A, AND 11A;		Query Match Score 98.3%; Best Local Similarity 98.5%; Pred. No. 5.4e-80;
RA		Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,		Matches 269; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
RA		"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the Inca 147T mutation.";		Qy 1 MTTPTLIVIPSPPPAPSYANRPQPSLMDKIKKTAIAASLILIGTIGFALLGHVGFL 60
RT		Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.		Db 1 MTTPTLIVIPSPPPAPSYANRPQPSLMDKIKKTAIAASLILIGTIGFALLGHVGFL 60
RL		DR EMBL; AF279346; AAC61099..-;		Qy 61 IAPQTIVVLLAFTSLAGNLYLQTANLHYLQDOREVGSKETINFMLSVLQKEFLHL 120
DR		DR EMBL; AF327329; AAK1123..-;		Db 61 IAPQTIVVLLAFTSLAGNLYLQTANLHYLQDOREVGSKETINFMLSVLQKEFLHL 120
DR		DR EMBL; AF326932; AAG61089..-;		Qy 121 SKEFATSKDLSAVSQDFYSCLOGFRDNYKGFESELDEYKNSDEEMRKLFSEBLIADIKG 180
DR		DR EMBL; AF326934; AAG61091..-;		Db 121 SKEFATSKDLSAVSQDFYSCLOGFRDNYKGFESELDEYKNSDEEMRKLFSEBLIADIKS 180
DR		DR EMBL; AF326938; AAG61095..-;		Qy 181 SVASLREIRFLPLAEEYRRLAHNOESTAAEELKLIRDSELDEIGOLSOLSKTISQ 240
DR		DR EMBL; AF327004; AAG61101..-;		Db 181 SVASLREIRFLPLAEEYRRLAHNRESLTAAEELKLIRDSELDEIGOLSOLSKTISQ 240
DR		DR EMBL; AF327006; AAG61103..-;		Qy 241 IALQRKESSDLCSQIRETLSSPRKSASSSTKSS 273
DR		DR EMBL; AF327011; AAG61109..-;		SEQUENCE 273 AA; 30327 MW; 6945AE7E8BBBEEF3 CRC64;
KW		Complete proteome.		
SQ		SEQUENCE 273 AA;		



				DE	INCLUSION MEMBRANE PROTEIN A.	
				GN	INCA.	
				OS	Chlamydia trachomatis.	
				OC	Bacteria; Chlamydiaceae; Chlamydia.	
				NCBI_TaxID	-813;	
				RN	[1]	
				RP	SEQUENCE FROM N.A.	
				RC	STRAIN=NT9329; AND MT9309;	
				RA	Viratoyin W., Rockey D.D., Suchland R.J., Stamm W.E.;	
				RT	"Diverse Mutations in inca Amplified from Clinical Chlamydia trachomatis Isolates that Occupy Non-fusogenic Inclusions.";	
				RL	Submitted (DCC-2000) to the EMBL/GenBank/DBJ databases.	
				DR	EMBL: AF327331; AAC11235.1;	
				DR	EMBL: AF327330; AAC11234.1;	
				SQ	SEQUENCE 273 AA; 30339 MW; B445B3638B8BEBE7 CRC64;	
RESULT 10						
Q9AMA7		PRELIMINARY;	PRP;	273 AA.		
ID Q9AMA7;						
AC Q9AMA7;						
DT 01-JUN-2001 (TREMBLrel. 17, Created)				97.9%	Score 1305; DB 2;	
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)				97.8%	Prod. No. 1.2e-79;	
DE INCLUSION MEMBRANE PROTEIN A.				Best Local Similarity	Mismatches 1; Indels 0; Gaps 0;	
OS Chlamydia trachomatis.				Matches 267; Conservative		
OC Bacteria; Chlamydiaceae; Chlamydia.						
OX NCBI_TaxID=813;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC STRAIN=DW-12931;						
RX MEDLINE=21295121; PubMed=11402010;						
RA Pannenkoek Y., van Der Ende A., Eijk P.P., van Marle J., de Witte M.A.,						
RA Ossewaarde J.M., van Den Brule A.J.C., Moire S.A., Dankert J.;						
RT "Normal inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the inca 147T mutation.";						
RT Infect. Immun. 69:4654-4656(2001).						
DR EMBL: AF327007; AAC61104.1; -;						
SQ SEQUENCE 273 AA; 30337 MW; 6945AE7E8BB8824F3 CRC64;						
Query Match 98.0%; Score 1306; DB 2; Length 273;				RESULT 12		
Best Local Similarity 97.8%; Pred. No. 9.9e-80;				Q9AMA8		
Matches 267; Conservative 1; Mismatches 5; Indels 0; Gaps 0;				AC Q9AMA8	PRELIMINARY; PRT; 273 AA.	
QY 1 MTPPTLIVPPSPAPPSYSANVPQPSLMDKIKKIAIASLILGTTGFLALLGHVGF 60				AC		
Db 1 MTPPTLIVPPSPAPPSYSANVPQPSLMDKIKKIAIASLILGTTGFLALLGHVGF 60				DT	(TREMBLrel. 17, Created)	
QY 61 IAPQITIVLALFITSLAGNALYLQQTANLHYQDLQREYGSLSKEINFMVLQEFHLH 120				DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	
Db 61 IAPQITIVLALFITSLAGNALYLQQTANLHYQDLQREYGSLSKEINFMVLQEFHLH 120				DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
QY 61 IAPQITIVLALFITSLAGNALYLQQTANLHYQDLQREYGSLSKEINFMVLQEFHLH 120				DE	INCLUSION MEMBRANE PROTEIN A.	
Db 61 IAPQITIVLALFITSLAGNALYLQQTANLHYQDLQREYGSLSKEINFMVLQEFHLH 120				GN	INCA.	
QY 121 SKEFATISKDLSAVSDFYSCLOQFRDNKGFEISLDEYKNSTEEMRKLESQEIADLKG 180				OS	Chlamydia trachomatis.	
Db 121 SKEFATISKDLSAVSDFYSCLOQFRDNKGFEISLDEYKNSTEEMRKLESQEIADLKG 180				OC	Bacteria; Chlamydiaceae; Chlamydia.	
QY 181 SVASLREERIRPLTPLAEEVRLLAHNGSLSLTAAIEEKTKTIRDSLDEIGQLSLSKTLTSO 240				NCBI_TaxID	-813;	
Db 181 SVASLREERIRPLTPLAEEVRLLAHNGSLSLTAAIEEKTKTIRDSLDEIGQLSLSKTLTSO 240				RN	[1]	
QY 241 IALQRRESDIQCQIRETLLSPRKSASPSSPKSS 273				RP	SEQUENCE FROM N.A.	
Db 241 IALQRRESDIQCQIRETLLSPRKSASPSSPKSS 273				RC	STRAIN=UW-12;	
QY 241 IALQRRESDIQCQIRETLLSPRKSASPSSPKSS 273				RX	Medline=21295121; PubMed=11402010;	
Db 241 IALQRRESDIQCQIRETLLSPRKSASPSSPKSS 273				RA	Pannekoek Y., van Der Ende A., Eijk P.P., van Marle J., de Witte M.A.,	
QY 241 IALQRRESDIQCQIRETLLSPRKSASPSSPKSS 273				RA	Ossewaarde J.M., van Den Brule A.J.C., Morre S.A., Dankert J.;	
Db 241 IALQRRESDIQCQIRETLLSPRKSASPSSPKSS 273				RT	"Normal inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the inca 147T mutation.";	
QY 241 IALQRRESDIQCQIRETLLSPRKSASPSSPKSS 273				RT	Infect. Immun. 69:4654-4656(2001).	
Db 241 IALQRRESDIQCQIRETLLSPRKSASPSSPKSS 273				RL	DR	EML: AF327005; AAC61102.1; -;
QY 241 IALQRRESDIQCQIRETLLSPRKSASPSSPKSS 273				DR	EMBL: AF327005; AAC61102.1; -;	
Db 241 IALQRRESDIQCQIRETLLSPRKSASPSSPKSS 273				SQ	SEQUENCE 273 AA; 30387 MW; 6945AE7E8BB8BEBF3 CRC64;	
RESULT 11						
Q9Q56		PRELIMINARY;	PRP;	273 AA.		
ID Q9Q56;						
AC Q9Q56;						
DT 01-JUN-2001 (TREMBLrel. 17, Created)				97.9%	Score 1305; DB 2;	
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)				97.8%	Prod. No. 1.2e-79;	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)				Best Local Similarity	Mismatches 1; Indels 0; Gaps 0;	
QY 1 MTPPTLIVPPSPAPPSYSANVPQPSLMDKIKKIAIASLILGTTGFLALLGHVGF 60				Matches 267; Conservative		

1	MTPPTLIVTPSPPPAPSYSANRVPQPSLMDKTKKIAIASLILITGFLALIGHYGFLL			
61	TAPQITIVLALFTISLAGNAYLQQTANHLHYQDQREVGSIKEINEMPLSVLQKEFHL	120		
61	TAPQITIVLALFTISLAGNAYLQQTANHLHYQDQREVGSIKEINEMPLSVLQKEFHL	120		
121	SKFEGATTSKDLSSAVSQDFYSCLOGFRDNKGFESELDEYKNSTBMRKLFSQEITADLK	180		
121	SKFEGATTSKDLSSAVSQDFYSCLOGFRDNKGFESELDEYKNSTBMRKLFSQEITADLK	180		
181	SVASLREEEIRFLTPLAEEVRLAHNEQESTAAEELKTIRDSLDEIGOLSQSKTLTSQ	240		
181	SVASLREEEIRFLTPLAEEVRLAHNEQESTAAEELKTIRDSLDEIGOLSQSKTLTSQ	240		
241	IAQLRKESSDLCSQIRETLLSPRKSASPSTKSS	273		
241	IAQLRKESSDLCSQIRETLLSPRKSASPSTKSS	273		
	SULT 13			
	BREX7	PRELIMINARY;	PRT;	273 AA.
	Q9RFX7			
	Q9RFX7;			
	C			
	C			
	01-MAY-2000 (TREMBLI)	13, Created)		
	01-MAY-2000 (TREMBLI)	13, Last sequence update)		
	01-JUN-2001 (TREMBLI)	17, Last annotation update)		
	INCLUSION MEMBRANE PROTEIN A.			
	INCA.			
	Chlamydia trachomatis.			
	Chlamydiales; Chlamydiaceae; Chlamydia.			
	[1]			
N	SEQUENCE FROM N.A.			
C	STRAIN=D(S)	2923;		
X	MEDLINE=2012/07/06;			
X	PubMed=10603409;			
X	Auchland R.J., Rooley D.D., Bannantine J.P., Stamm W.E.;			
X	"Isolates of Chlamydia trachomatis that occupy nonfusogenic inclusions lack incA, a protein localized to the inclusion membrane.";			
L	Infect. Immun. 68:3650-3671(2000).			

```

Query Match      Score 1302; DB 2; Length 273;
Best Local Similarity 97.7%; Pred. No. 1.8e-79;
Matches 266; Conservative 2; Mismatches 5; Indels 0; Gaps 0
1 MTTPTLIVIPTPPSPAPSYSANVPQPSLMDKIKKIAAIASLLILIGTGFLLAIGHLYGFL 60
1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||| |
1 MTTPTLIVTPTPPSPAPSYSANVPQPSLMDKIKKIAAIASLLILIGTGFLLAIGHLYGFL 60
1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||| |
61 IAPQTIVLLAFLITSAGNAYLQKTANLHYQLOREVGSIKEINFMLSVIQLKEFLHHL 120

```

DE INCLUSION MEMBRANE PROTEIN A.  
 GN INCA.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydaceae; Chlamydia.  
 OX NCBI\_TAXID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UW-1;  
 RX MEDLINE=21205121; PubMed=11402010;  
 RA Pannenkoek Y., van Der Ende A., Eijk P.P., van Marle J., de Witte M.A.,  
 RA Ossenwaarde J.M., van Den Brule A.J.C., Morre S.A., Dankert J.;  
 RT "Normal Inca expression and fusogenicity of inclusions in Chlamydia  
 RT trachomatis isolates with the inca I47T mutation.",  
 RL Infect. Immun. 69:4654-4656 (2001).  
 DR EMBL: AF325995; AAG61092.; -  
 SQ SEQUENCE 273 AA; 30364 MW; 0549E32D7979B8A CRC64;

Query Match 96.9%; Score 1292; DB 2; Length 273;  
 Best Local Similarity 97.4%; Pred. No. 8.5-e-79;  
 Matches 266; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 / 1 MTTPPLIVIPPPPPAPSYSANRVPQPSMIDKIKKIAATAASLLIGTGFALLGLYFL 60  
 Db 1 MTTPPLIVTPSPPPAPSYSANRVPQPSMIDKIKKIAATAASLLIGTGFALLGLYFL 60  
 Qy 61 IAQITIVYLALFTISLAGNAYLQKTNHLHQDLOREVGSLSKEINPMLSVLQKEFLHL 120  
 Db 61 IAQITIVYLALFTISLAGNAYLQKTNPHLYQDLOREVGSLSKEINPMLSVLQKEFLHL 120  
 Qy 121 SKEFATTSKDLSAVSDQFYSCLOGFRDNKGPFESLIDYKNSTEEMRLFSQEIIADLG 180  
 Db 121 SKEFATTSKDLSAVSDQFYSCLOGFRDNKGPFESLIDYKNSTEEMRLFSQEIIADLG 180  
 Qy 181 SVASLREERFLTPLAEVRLAHHNQELDTAAIEELKTIRDSLDEIGQLSLSKLTLSQ 240  
 Db 181 SVASLREERFLTPLAEVRLAHHNQSLTVVIEELKTIRDSLDEIGQLSLSKLTLSQ 240  
 Qy 241 IALQRKESSDLCSQIRETLLSPRSKASPTKSS 273  
 Db 241 IALQRKESSDLCSQIRETLLSPRSKASPTKSS 273

Search completed: August 13, 2002, 09:30:17  
 Job time: 373 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: August 13, 2002, 09:22:24 ; Search time 29.79 Seconds  
(without alignments)  
1145.073 Million cell updates/sec

Title: US-09-673-763-8  
Perfect score: 1733  
Sequence: 1 MTVDNTSPVIRSASSPTF.....DDSAQEQDENQSAGEHKDS 355

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PUR\_71;\*  
1: Pir1;\*  
2: Pir2;\*  
3: Pir3;\*  
4: Pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1733	100.0	355	2 S61491	inclusion membrane hypothetical prote
2	200	11.5	390	2 A72108	hypothetical prote
3	200	11.5	390	2 BB6514	probable inclusion membrane
4	185	10.7	276	2 H8107	hypothetical prote
5	164.5	9.5	273	2 C71553	probable inclusion membrane
6	149	8.6	533	2 G72593	hypothetical prote
7	145.5	8.4	1732	2 T14039	protein kinase (EC
8	145	8.4	1690	2 T13030	microtubule bindin
9	143	8.3	1410	1 A57013	early endosome ant
10	141	8.1	944	2 S26710	spindle pole body
11	140.5	8.1	403	2 H72216	outer membrane pro
12	140	8.1	770	2 F82383	methyl-accepting C
13	140	8.1	1056	1 G02157	kinesin-like spin
14	140	8.1	1155	2 AC275	chromosome segreg
15	140	8.1	1165	2 A97457	structural mainten
16	139.5	8.0	739	2 H75001	methyl-accepting C
17	139.5	8.0	746	2 T47237	myosin II heavy ch
18	139	8.0	1805	1 A61224	hypothetical prote
19	138.5	8.0	772	2 T22907	ciltron - mouse
20	138.5	8.0	1597	2 S68420	skeletal myosin -
21	138.5	8.0	1957	2 A59294	Paranyosin - fluke
22	137	7.9	439	2 A60608	hypothetical prote
23	136.5	7.9	1133	2 T22976	myosin in heavy chain
24	136.5	7.9	1940	1 S04090	male-enhanced anti
25	136.5	7.9	1325	1 A24922	apsB protein - Eme
26	136	7.8	1790	2 T47222	myosin heavy chain
27	135.5	7.8	1051	2 T18302	interaptin - slime
28	135.5	7.8	1937	2 T38055	Chlamydophila pneumoniae (strains CWL029 and
29	135	7.8	1738	2 T14867	C; Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

Result	Match	Length	DB ID	Score	Local Similarity	Mismatches	Indels	Gaps
1	1	1	S61491	100.0%	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Qy	1	1	MTVSTDNTPSPVIRSASSPTF	100.0%	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Db	1	1	MTVSTDNTPSPVIRSASSPTF	100.0%	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Qy	61	1	LQRICLYKLIAATLFYVGIAALVCLYLGSVISTPSLTLMLATMLVSFIVITARDG	120	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Db	61	1	LQRICLYKLIAATLFYVGIAALVCLYLGSVISTPSLTLMLATMLVSFIVITARDG	120	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Qy	121	1	TPSQVRHMKQIQQGEENTRLAVENLRKAVNVELQAVNLTRLSDFGDRLEA	180	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Db	121	1	TPSQVRHMKQIQQGEENTRLAVENLRKAVNVELQAVNLTRLSDFGDRLEA	180	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Qy	181	1	NTGDFIALIADEQLSLEEFKSVGTKVETMSPFKELAQSKETEVSQEAQAMSSVTTEL	240	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Db	181	1	NTGDFIALIADEQLSLEEFKSVGTKVETMSPFKELAQSKETEVSQEAQAMSSVTTEL	240	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Qy	241	1	TNLNAKELLTENKTYIEOLKADAQLEQEVRFLEKRGKQEEACSTLSHSTATLQESTT	300	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Db	241	1	TNLNAKELLTENKTYIEOLKADAQLEQEVRFLEKRGKQEEACSTLSHSTATLQESTT	300	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Qy	301	1	LLKDSTNLHAYESRLIGTMVQDGAESSTVEASQDSDAQPDENQSDAGEHKDS	355	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
Db	301	1	LLKDSTNLHAYESRLIGTMVQDGAESSTVEASQDSDAQPDENQSDAGEHKDS	355	Score 1733;	DB 2;	Length 355;	
					Best Local Matches	100.0%;	Freq. No. 4.3e-93;	
					Conservative	0;	Mismatches 0;	
							Indels 0;	
							Gaps 0;	
RESULT	2							

A72108 hypothetical protein CP058 [Imported] - Chlamydophila pneumoniae (strains CWL029 and

C; Species: Chlamydophila pneumoniae, Chlamydia pneumoniae, Chlamydia pneumoniae

C; Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000

C; Accession: A72108; B81560

R; Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;	Query Match Score 11.5%; Best Local Similarity 23.0%; Pred. No. 0.00015;
A; Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.	Best Local Similarity 23.0%; Mismatches 64; Indels 86; Gaps 13;
A; Reference number: A72000; MUID:99206606	Matches 87; Conservative 64; Mismatches 142;
A; Accession: A72108	
A; Molecule type: DNA	
A; Cross-references: GB:AE001605; GB:AB001363; NID:94376455; PIDN:AAD18339.1; PID:9437645	Qy 8 TSPVIRASSPFFGDKDFDNKKLIPISIAPTSAAAVGAKTAALEPEGRSPILLORICY 67
A; Experimental source: strain CWL029	Db 2 SSPVNNTPSAPN-----IPIPAPTP---GIPPT--KPKSSFIEKVI- 38
A; Residues: 1-390 <ARN>	Qy 68 LYKIIAIAALFYVGIALVCLYLGSVI-S---TPSL-ILMLAIMLYSFIVITAIIRDGT 121
A; Read, T.D.; Brumham, R.C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000	Db 39 --IVAKYIFAI---ATSGALTIGLGSALTPGIGIALVIVFFMSVLGLILRDSI 92
A; Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.	Db 122 PSQYRHMKQOIQQFGBEN-----TRLHTAVENLKAVN----- 154
A; Reference number: A81500; MUID:20150255	Qy 1 SGGERRRLREEVSRFTSENQRQTVITTEVKDLRAKDQTLTEAFRNENGNLKTT 152
A; Molecule type: DNA	Db 93 LREEQVRFLERKQEEACSTLHSIATLOESTTLLKDTSTNLHAYEVQYDGA 212
A; Residues: 1-390 <REA>	Qy 155 -YPLSEQINQKLQHTRLSDFGDRLEANTGDTALIAFDQLSLEEFNS-VGTKVETMSP 212
A; Cross-references: GB:AE002217; GB:AB002161; NID:g7189493; PIDN:AAF38395.1; PID:9718949	Db 153 AEDIEEQQVSKLSBQLEALERINOLIQANAGDAQEISSELKLIKISWMSKVYEQINTSI-- 210
A; Experimental source: strain AR39, HL cells	Qy 213 FEKLAQSLKETFSQEAQYQAMMSVYTELRTNLNAKELI-----TENKTVIEOLKAQAO 265
A; Genetics:	Db 211 ----OALKVILQGQEWVQDQWAAQTVIKAMQEQIQLQATELGMNQSTALQKSVENLLVQDQ 265
A; Gene: Cpn0186; CP0581	Qy 266 LREEQVRFLERKQEEACSTLHSIATLOESTTLLKDTSTNLHAYEVQYDGA 325
RESULT 4	Db 266 ALTRVYGEELLESENKLQSACLSQASLRQEIEKLAQHETSLSQRIDAMLAOEONL-----A 317
Qy 68 LYKIIAIAALFYVGIALVCLYLGSVI-S---TPSL-ILMLAIMLYSFIVITAIIRDGT 121	Qy 326 ESSTVEEASQDDSAQPODE 344
Db 39 --IVAKYIFAI---ATSGALTIGLGSALTPGIGIALVIVFFMSVLGLILRDSI 92	Db 318 EQVTALEKMKQEAQAESE 336
Qy 122 PSQYRHMKQOIQQFGBEN-----TRLHTAVENLKAVN----- 154	RESULT 5
Db 93 SGGERRRLREEVSRFTSENQRQTVITTEVKDLRAKDQTLTEAFRNENGNLKTT 152	H81/07 inclusion membrane localised protein Inca TC0396 [Imported MoPn] - Chlamydia muridarum (strain H81/07)
Qy 155 -YPLSEQINQKLQHTRLSDFGDRLEANTGDTALIAFDQLSLEEFNS-VGTKVETMSP 212	C; Species: Chlamydia muridarum, Chlamydia trachomatis
Db 153 AEDIEEQQVSKLSBQLEALERINOLIQANAGDAQEISSELKLIKISWMSKVYEQINTSI-- 210	C; Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
Qy 213 FEKLAQSLKETFSQEAQYQAMMSVYTELRTNLNAKELI-----TENKTVIEOLKAQAO 265	C; Accession: H81107
Db 211 ----OALKVILQGQEWVQDQWAAQTVIKAMQEQIQLQATELGMNQSTALQKSVENLLVQDQ 265	C; Read, T.D.; Brumham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
Qy 266 LREEQVRFLERKQEEACSTLHSIATLOESTTLLKDTSTNLHAYEVQYDGA 325	A; Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
Db 266 ALTRVYGEELLESENKLQSACLSQASLRQEIEKLAQHETSLSQRIDAMLAOEONL-----A 317	A; Reference number: A81500; MUID:20150255
Qy 326 ESSTVEEASQDDSAQPODE 344	A; Accession: H81107
Db 318 EQVTALEKMKQEAQAESE 336	A; Status: preliminary
RESULT 3	A; Molecule type: DNA
B8514 hypothetical protein CPJ0186 [Imported] - Chlamydophila pneumoniae (strain J138)	A; Residues: 1-776 <TET>
C; Species: Chlamydophila pneumoniae, Chlamydia pneumoniae	A; Cross-references: GB:AE002306; GB:AE002160; PIDN:9719042B; PIDN:AAF39253.1; PID:9719
C; Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001	A; Genetics:
C; Accession: B86514	A; Gene: TC0396
R; Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; IS Nucleic Acids Res. 28, 2311-2314, 2000	Query Match Score 10.7%; Best Local Similarity 23.8%; Pred. No. 0.00075;
A; Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.	Matches 71; Conservative 64; Mismatches 121; Indels 48; Gaps 11;
A; Reference number: A86491; MUID:20330349	Qy 16 SSPTFDGHGKDFDNKKLIPISIAP-TSSAAVGAKTAALEPEGRSPILLORICYLVKIIIA 73
A; Status: preliminary	Db 2 TSPTLVE-----MPLSCYPHTTSACTKRSSLYKPSLITERVQAAPFSL-- 49
A; Molecule type: DNA	Qy 74 AIALFYVGIAAVCLYLGSVI-S---TPSL-ILMLAIMLYSFIVITAIIRDGTSPSVVRHKQOI 133
A; Residues: 1-390 <STO>	Db 50 -AILSIGFLAILGHAGIELFAPQVAVLTAFTISLLGNALYLCATPLRYKEIQQEV 108
A; Cross-references: GB:BA000008; NID:98978559; PIDN:BAA98396.1; GSPDB:GN00142	Qy 134 QQFGENTRLHTAVENLKAVNVELSBOINQHLRSLDEGDRLEANTGDTALIAFDQ 193
C; Genetics:	Db 109 ASLKEYNFL-----LKSVOKEF----VSLDFHNLQDFD 154
A; Gene: CPJ0186	Qy 194 LSLEEFKSVGTKVETMLSPFKELAQSLKETFSQEAQYQAMMSVTEIRTNLNALKEI----- 249

Db 155 SSHQGF-----EDLDIYDKNSAEDLQIFSQETVQSLKSTLSEKEIKEIVPLDEEV 207  
 A; Experimental source: strain K1  
 C; Genetics:  
 A; Gene: APE1216

Qy 250 --TRENKTYIEQLKADAQ-LREEQVRFELRKQDELEEACSTLHSIA-TLOESTTLIKDS 305  
 Db 208 RRAENKDLLKTVQDLDQDIRK---LRAEINNLQSASKLSEQASQEENEKLYANI 263  
 Qy 306 TTNL 309  
 Db 264 TKAL 267

RESULT 5  
 C71553 probable inclusion membrane protein A - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C;Species: Chlamydia trachomatis  
 C;Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
 C;Accession: C71553  
 R; Stephens, R.S.; Kalmel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
 A; Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac Accession: C71553  
 A; Status: Preliminary  
 A; Molecule type: DNA  
 A; Residues: 1-273 <ARN>  
 A; Cross-references: GB:AE001286; GB:AE001273; NID:93328516; PID:AACC67710.1; PID:g332851  
 C; Genetics:  
 A; Gene: incA

Query Match 9.5%; Score 164.5; DB 2; Length 273;  
 Best Local Similarity 22.8%; Pred. No. 0.011;  
 Matches 67; Conservative 59; Mismatches 109; Indels 59; Gaps 10;  
 Qy 32 LIPISTEAPNNSAAAVGAKATAIEPPGRSPSLLQRCYLVIIIAIALFYVGAALVCLYLG 91  
 Db 8 VTPPSPPAPSYSANRV-----POPSLMDKIKKIAIASLILIGTGFALLGLHVG 58  
 Qy 92 SVISPSPLIMLAIMLVSVIVITAIRDTPSQVVR--HMKQOIQIOFGEENTRLHTAV 147  
 Db 59 FLIAQQTIVLLALFTIISLA-----GNAVYIQLKTANLHLTDLQR-----EV 100  
 Qy 148 ENLAKVNVEUQELSEQNOLQKOLHTRLSDFGDRLEANTGDTALIADFLSLEEFSVGTKE 207  
 Db 101 GSUREINFLS--VLOQKELH----LSKEFATTSKDLSAVSDFYSCLOGFRDNKGFE 153  
 Qy 208 TMSPFEKAQSLKETFSQEAQAMMSSVTELRTNLNALKELENK-----TVI 257  
 154 SLDEYKNSTEEMRKLFQSEITIADLKGSVASLREEIRPLAAEHNAHQSLTVVI 213  
 Qy 258 EQIKA-DAQLEE-QVRFLEK-----RQEQLERACTLHSIA-TLOESTT 300  
 Db 214 EELKTIRDSSLRDEIGLQSQSLKTLTSQIALQRKESSDCISQIRETLLSPRSAS 267

RESULT 6  
 G72593 hypothetical protein APE1216 - Aeropyrum pernix (strain K1)  
 C;Species: Aeropyrum pernix  
 C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
 C;Accession: G72593  
 R; Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahava, H.; Tatsunagi, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999  
 A; Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr A; Accession: G72593  
 A; Status: Preliminary  
 A; Molecule type: DNA  
 A; Residues: 1-533 <ARN>  
 A; Cross-references: DDBJ:AP000061; NID:95104821; PID:1043991; PID:g510

Query Match 8.6%; Score 149; DB 2; Length 533;  
 Best Local Similarity 24.0%; Pred. No. 0.2;  
 Matches 59; Conservative 46; Mismatches 109; Indels 32; Gaps 7;  
 Qy 124 QVVRHMKQQTOFGENTRHATVENVLNKAVNVELQEQINOLQKOLHTRLSDFGDRLEANTG 183  
 Db 251 QAVEDIGLADSLEERVGDLFEAVEDL---TLQLSLSDSERVGALEDRVADIEGREAVEG 307  
 Qy 184 DF-----TALIADFLSBEFKSVGTKVETMLSPFEKLAQSLKETFSQ--EAVQAMHS 234  
 Db 308 SLEDLSGAVDAMSQQLQALADELSSLESSRYEDL---EARVGSYEDRUSQAEEEDDSLT 363  
 Qy 235 SVTERTNNALKELENKTVIEQLKADAQRLREEVRFLEKREKQELBEACSTLHSIA-T 294  
 Db 364 SLDSLRLTEEDLSTLIAEAQSLEDLNTRLQVASTLQQLQRLATAEBSLQLATEQDIA 423  
 Qy 295 LQESTTLLKDSTTNHAYESRLTIGMVQDGAEESTV-----EASQDDSAQPQDENOSD 348  
 Db 424 LQAEVETLQOSIVE---IDRRL-----GQLRSVTDAVLREYESLGKLYQAEEKNRQ 473  
 Qy 349 AGEHRD 354  
 Db 474 DASTBD 479

RESULT 7  
 T14039 protein kinase (EC 2.7.1.37), myotonic dystrophy-associated - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jun-2000  
 C;Accession: T14039  
 R; Leung, T.; Chen, X.O.; Tan, I.; Manser, E.; Lim, L.  
 Mol. Cell. Biol. 18, 130-140, 1998  
 A; Title: Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a Cdc42 effec A; Reference number: Z17862; MUID:98078670  
 A; Status: preliminary  
 A; Molecule type: mRNA  
 A; Residues: 1-1732 <LEU>  
 A; Cross-references: EMBL:AE021935; NID:92736151; PID:92736151; PID: AAC02941.1  
 C; Genetics:  
 A; Gene: MRCK  
 C; Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog F; 75-343/Domain: protein kinase homology <KIN>  
 F; 1013-1062/Domain: protein kinase C zinc-binding repeat homology <KZND>  
 Query Match 8.4%; Score 145.5; DB 2; Length 1732;  
 Best Local Similarity 21.1%; Pred. No. 1.3;  
 Matches 54; Conservative 59; Mismatches 114; Indels 29; Gaps 7;  
 Qy 126 VRHMKQQIOQFGENTRHATVENVLNKAVNVELQEQINOLQKOLHTRLSDFGDRLEANTGDF 185  
 Db 506 VNHLSQLQEAEANSVRELLDAFRQIKAFEQKIKLQLQEREEELNLKQSKEL 565  
 Qy 186 TALIADFLSBEFKSVGTKVETMLSPFEKLAQSLKETFSQAVQAMMSVTELRNL-- 243  
 Db 566 KDAHQQRKLAMQEPMEINERLTDHTQKLARHYD--KEEYDVQKAEQSRQELRR 623  
 Qy 244 --NALKELLENKTVIEQLKADAQRLREEVRFLEKREKQELBEACSTLHSIA-T 295  
 Db 624 AERAKTLEVHTTEALIAEASKRKLREQSRHYSKQLENELEGLKQKQTSYSPGICSHQ 683  
 Qy 296 QESTKLLKD-----STNLIAEVASRIGVMQD--GAESTVVE--ASODDS 338  
 Db 684 QETTKLTKDLEKKSIFFEEEISKRGGIAEIRNLKKEQQLALNKEIMVLDKL 743

QY 339 AQPDENOSDAGEHKD 354  
 A; Reference number: S44243  
 A; Accession: S44243  
 A; Status: preliminary  
 A; Molecule type: mRNA  
 A; Residues: 1-54; 'C', 256-257, 'LQ', 260-276, 'A', 278-283, 'A', 285-519, 'D', 521-574, 'EQ', 5.  
 C; Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
 C; Species: Drosophila melanogaster  
 C; Accession: T13030  
 A; Reference number: 217588; MUID:98139549  
 A; Status: preliminary; translated from GB/EMBL/DDJB  
 A; Molecule type: mRNA  
 A; Residues: 1-1690 <LAN>  
 Cross-references: EMBL:AF041382; NID:g2773362; PID:92773363; PID:AA96783.1  
 A; Experimental source: strain Oregon R  
 C; Genetics:  
 A; Cross-references: FlyBase:FBgn0020503  
 C; Keywords: cytoskeleton

RESULT 8  
 T13030 microtubule binding protein D-CLIP-190 - fruit fly (Drosophila melanogaster)  
 C;Species: Drosophila melanogaster  
 C;Accession: T13030  
 A;Title: A class VI unconventional myosin is associated with a homologue of a microtubul  
 J. Cell Biol. 140, 97-910, 1998  
 A;Reference number: 217588; MUID:98139549  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: mRNA  
 A;Residues: 1-1690 <LAN>  
 Cross-references: EMBL:AF041382; NID:g2773362; PID:92773363; PID:AA96783.1  
 A;Experimental source: strain Oregon R  
 C;Genetics:  
 A;Cross-references: FlyBase:FBgn0020503  
 C;Keywords: cytoskeleton

Query Match 8.4%; Score 143; DB 1; Length 1410;  
 Best Local Similarity 24.0%; Pred. No. 1.4;  
 Matches 64; Conservative 52; Gaps 11;  
 Query Match 8.3%; Score 143; DB 1; Length 1410;  
 Best Local Similarity 24.0%; Pred. No. 1.4;  
 Matches 64; Conservative 52; Gaps 11;  
 QY 129 MKQQIOPGEE -----NTRLHTAVENLKVNVTESEQINOLKQHTRLSDFGDRLEA 180  
 Db 812 LKQDFETLSQETKIQHEELNNRIOCTTTELQVKMKPEALMTSTVTKDLSKVNQSDSLKN 871  
 QY 181 NTGDF-----TALADFOISLEEK-SVGTQKVEMLMSPEKKLAQS-LKETFSQEAYOA 231  
 Db 872 SKSEFEKENQRGKAAILDLKETCKELKHLQHQVOMENTLKKEELKSLEKEASHQKL 931  
 QY 232 MHSVTE-----LTINLNALKLTENKTVIQOLKADAQQLREEQVRF 273  
 Db 932 ELNSMQEFLQIAQNTLKONEKEQQLQGNNINLKQSEQQKQEAQGEKLTIAVQLOTE 991  
 QY 274 LE-KRKQLEEEACSTS--HSTATLQ-----ESTLRLDSTTNLHAVESRLIGVMVQDG 325  
 Db 992 LENLQDQOITQAQELAAKEKTSVQLYNNYERKSQETFKQLQSDFYGRGESELLATR---Q 1047  
 QY 326 ESSTVEE--ASQDD--SAQPODENOS 347  
 Db 1048 DLKSVEEKLSSLAQEDLISRNQIGNQN 1074  
 RESULT 10  
 S26710 NUF1 - yeast (Saccharomyces cerevisiae)  
 spindle pole body protein NUF1 - yeast (Saccharomyces cerevisiae)  
 N;Alternate names: protein D9476; 3'; protein YDR356w  
 C;Species: Saccharomyces cerevisiae  
 C;Date: 07-May-1993 #sequence\_change 07-May-1993 #text\_change 23-Mar-2001  
 C;Accession: S26710; S34288; A49455; S61152  
 R;Mirzayan, C.; Copeland, C.S.; Snyder, M.  
 J. Cell Biol. 116, 1319-1332, 1992  
 A;Title: The NUF1 gene encodes an essential coiled-coil related protein that is a pot  
 A;Reference number: S26710; MUID:9217632  
 A;Accession: S26710  
 A;Molecule type: DNA  
 A;Residues: 1-944 <MIR>  
 A;Cross-references: EMBL:11562; NID:94069; PID:CAA77668.1; PID:94070  
 A;Note: the authors translated the codon GAG for residue 206 as AsP and CTG for resid  
 R;Kilmartin, J.W.; Dyos, S.L.; Kershaw, D.; Finch, J.T.  
 submitted to the EMBL Data Library, June 1993  
 A;Description: A spacer element in the Saccharomyces cerevisiae spindle pole body who  
 A;Reference number: S34287  
 A;Accession: S34288  
 A;Molecule type: DNA  
 A;Residues: 1-944 <KII>  
 A;Cross-references: EMBL:X73297; NID:g312173; PID:9312175  
 B;Kilmartin, J.W.; Dyos, S.L.; Kershaw, D.; Finch, J.T.  
 J. Cell Biol. 123, 1175-1184, 1993  
 A;Title: A spacer protein in the Saccharomyces cerevisiae spindle pole body whose tra  
 A;Reference number: A49455  
 A;Accession: A49455  
 A;Molecule type: DNA  
 A;Residues: 1-22 <K12>  
 A;Cross-references: EMBL:X73297  
 R;Du, Z.  
 submitted to the EMBL Data Library, June 1995  
 A;Description: The sequence of S. cerevisiae cosmid 9476.  
 R;Seelig, H.P.  
 submitted to the EMBL Data Library, April 1994

A; Reference number:	S61148	Db	146 QSDEQLSLDAKVNEALSKIAALESKLGLGDFVNQDVKVSKLSDLEGRLSVAETK 205
A; Accession number:	S61152	Qy	182 TGDFFALIADFQLSLEEFKSVGTVKVTMSPFKEKLAQSLKETFSQEAQYAMMSVTELR 241
A; Molecule type: DNA		Db	206 TANLALVNRSEASIKDY -----VENTLKSYTDTDLQKLSLSAS---VEK 248
A; Residues: 1-94 <DNU>		Qy	242 NLNAALKELNKTENKTYIEQLKADAQLEEVRFLEKROEELACSTLSHSTATLQESTT 301
A; Cross-references: SGD:NUF1; SPC110		Db	249 NNTAUSGETGNLNLYVLSKQLQSDLTETQQKTTARALDARYSVLGQTIVNSRVELEKRVSQ 308
A; Gene: SGD:NUF1;		Qy	302 LKDSTNLHAYESPLIGYVWQDAAESTVEASQDDSAQPDENQSDA 349
A; Cross-references: EMBL:U28372; NID:9849173; MIPS:YDR356w		Db	309 VESAVDKVNSLE-RSMGAVT---ARVTKVEE-----EVKNLQNSNA 345
A; Map position: 4R			
C; Keywords: coiled coil; microtubule; nucleus			
Query Match	8.1%; Score 141; DB 2; Length 944;		
Best Local Similarity	23.5%; Pred. No. 1.1;		
Matches 61;	Conservative 56; Mismatches 101; Indels 42; Gaps 8;		
Qy	129 MKQQTQOFGEBNTRLHTAVENL-----AVNVLESEQINQLQKHLTRLSDFGDR 177	RESULT 1.2	
Db	316 MDLQLKQKONESKRLKDENLEKTFSENGQSAAKENLMLRKNIKFLLEEEITSKNSQ 375	F82383	methyl-accepting chemotaxis protein VCA1056 [Imported] - Vibrio cholerae (strain NJ59)
Qy	178 LEANTGDFTAFLADFQLSLEEFK-----SVGTKVETMSPFEKLAQSLK---ETFSQ- 226	C; Species: Vibrio cholerae	
D.	376 LIAKEGKKLASLMA -QLTQLESKLQRDSQLGSREEELKTNKLQNDRIAREETISKD 433	C; Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001	
Qy	227 EAVQAMMSSVTELTNLNLKELTENKTV-----IEQLKADAQLEEVRFLE 275	C; Accession: F82383	
D.	434 ERIDLQKVKVQLDLPVTKTHSEKTTDNELESKDKLILNDLVKAQEKKYSME 493	R; Heidelberg, J.F.; Eisen, J.A.; Neilson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.; Richardson, D.; Erauilaeva, M.D.; Yamatievean, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellest, L.R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.	
Qy	276 KROELEECACSTLSHSTATLQESTTLLKSTTNLHAYESPLIGYVWQDAAESTVEEAQ 335	A; Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.	
D.	494 K---ELKEEREFNYKISEKLEDEKTTLNKEKISNIAENSQLRNKI---EDNSTATHMK 546	A; Reference number: A82035; MUID:20406833	
Qy	336 DDSAQPDENQDAGEHKDS 355	A; Status: preliminary	
D.	547 ENYEKOLESLRKDIEEYKES 566	A; Molecule type: DNA	
		A; Residues: 1-770 <HEI>	
		A; Cross-references: GB:AE004432; GB:AE003853; NID:99658509; PIDN:AAF96950.1; GSPDB:EN	
		A; Experimental source: serogroup O1; strain N16961; biotype El Tor	
		C; Genetics:	
		A; Gene: VCA1056	
		A; Map position: 2	
RESULT 1.1			
H72216	outer membrane protein - Thermotoga maritima (strain MSBB)		
C; Species: Thermotoga maritima		Query Match	8.1%; Score 140; DB 2; Length 770;
C; Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000		Best Local Similarity	21.6%; Pred. No. 1;
C; Accession: H72216		Matches 73; Conservative 60; Mismatches 139; Indels 66; Gaps 13;	
R; Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, C.M.		Qy	68 LVKTIATALEFVGVIAFLYCLGVSIVST--SLIMLAIMLVSFVIVITAIRDGTP--- 122
Nature 399, 323-329, 1999		Db	417 IMKVLIATIVIGAV-ITALAAWMVNGISKPIRDTSIQFMRNDLTVRLSEOGSDIEIR 475
A; Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence		Qy	123 -----SQYVRHMKQIOQGEE---NTRLHTAVENLKVAVNYLESEQ-----IN 162
A; Reference number: A72200; MUID:99887316		Db	476 QLAQALNRMISHQDQTICQAHTDKLNSHTQTITHNPTGTRNSVSEHOERTDSVVATVN 535
A; Molecule type: DNA		Qy	163 QLKQHTRLSDFGDRL----EANTGDFEFALIAFDQLSLEEFKSVGTKVETMSPFPELA 217
A; Residues: 1-403 <ARN>		Db	536 EMSASITVESEFAORATPVQEANQKHHGVSGVNELA-RDMTSIN---QMASAVEAIA 591
A; Cross-references: GB:AE001812; GB:AE000512; NID:94982302; PIDN:AAD36794.1; PID:g498230		Qy	218 QSLKETFSQEAQYAMMSVTELPNLNLKELIENKT-----VTEQLKADAQ 265
A; Experimental source: strain MSB8		Db	592 RLNFHQSQTASVLDQIAE-OTNLALNAAEAARAGEQGRGFAYADEFRNLAAKTQ 650
C; Genetics:		Qy	266 LREEQVRF---LEKRKQE---LEEAGSTLSHSTATLQESTTLLKSDTNNLHAEVSLR 317
A; Gene: TM1729		Db	651 TSTBEIRTKIDRQKETOSVNVTCIEANNTVYRGVACHSNNTDMLKQIVDMUNELNEMNI 710
RESULT 1.3		Qy	318 -----GVWVQDGAEESTVEASQDDSAQPDEN 345
G02157		Db	711 QIATATEQORGVTEEINNNTTSVDASVTVQGEVN 748
kinesin-like spindle protein HKSP - human		Qy	
C; Species: Homo sapiens (man)			
C; Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001			
C; accession: G02157			
Query Match	8.1%; Score 140.5; DB 2; Length 403;		
Best Local Similarity	21.0%; Pred. No. 0.44;		
Matches 73; Conservative 75; Mismatches 113; Indels 87; Gaps 15;			
Qy	47 VGAKTAIEPGRSPSLLPQRLCYLVKLTIAIAFLVYGAIAVLVCLYLGSVISTPSLILMLAIM 106		
D.	40 LGVLIGM-PHG---TFOGNSYLTRYQRAYFY-----RLYNILKOPSDAVSGLIN 85		
Qy	107 LVSFV -IVTAIR-----DGTGPSY-----VRYHMK-----QQIQQ 135		
D.	86 KVSTLEDLYSTALMKVQNLSDNFGVTSDLTELKNDVANLAKTLVDLKNRVEVMSQVQS 145		
Qy	136 FGEE ---NTRLHTAVENLKVAVNYLESEQ-----NQLQHTRLSDFGDRL---EAN 181		

R;Whitehead, C.;	Submitted to the EMBL Data Library, September 1995
A;Reference number: H00839	
A;Accession: G02157	
A;Status: preliminary; translated from GB/EMBL/DDBJ	
A;Molecule type: mRNA	
A;Residues: 1-1056 <WHI>	
A;Cross-references: EMBL:U37426; NID:91171152; PID:AAA86132.1; PID:91171153	
C;Genetics:	
A;Gene: GDB:KNSL1; Eg5; KSP	
A;Cross references: GDB:132856; OMIM:148760	
A;Map position: 16q24.1-10q22.1	
C;Superfamily: kinesin-related protein Eg5; kinesin motor domain homology	
C;Keywords: APP; nucleotide binding; P-loop	
F;19-365/Domain: kinesin motor domain homology <KMOT>	
F;105-112/Region: nucleotide-binding motif A (P-loop)	
every Match 8.1%; Score 140; DB 1; Length 1056;	
Best Local Similarity 21.7%; Pred. No. 1.5;	
Matches 56; Conservatve 56; Mismatches 100; Indels 46; Gaps 10	
Qy 113 VITAIRDGT-----SQVYRHMKCOIQOFGENTRLHAVENLKAVNVELSEQINQKQ 166	
Dy 298 VITALVERTHPVYTESKLTRILQSGL-G-GRTRTSI---IATISPASNLNEELTSLLEY 352	
Qy 167 LHTRLSDFGDRLEANTG-DFTALIAFQLSLEEFK---SVGTKVETMLSPFEKLAQSLK 221	
Dy 353 AH-RAKNKLNRPEVNQKLTKRALKIKEYTEILERKRDADAREKNGVYISEENFRVMGK 411	
Qy 222 ETFSQEAVQAMMSSVTELRNLNALKELTENKTYEQLQKADAQREQTFLERKQKL 281	
Dy 412 LTVQEQUIVELETIKGAVEEEPLRNVLFMNKNEEDQCKSLDNQNTQELTTQKHQET 471	
Qy 282 BEACSTLSHSTATLQES--TFLKDSTNLHAYESRLLGYMVQDGAEASSTVEEAQSD--- 336	
Dy 472 K-----LQLVKBEYITSALESTEKLJHDAAKSL-----NFTVEETTRDVG 512	
Qy 337 -----DQAQPQDENOSDA 349	
Dy 513 LHSKLDKKAVDOHNAEA 530	
RESULT 14	
AC2675 chromosome segregation protein [imported] - Agrobacterium tumefaciens (strain	
Species: Agrobacterium tumefaciens	
Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002	
Accession: AC2675	
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, G.; Gillett, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, Y.; Karp, P.; Romero, P.; Zhang, S.	
Science 294, 2317-2323, 2001.	
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Goster, E.W.	
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens	
A;Reference number: AB2577; PMID:11743193	
A;Accession: AC2675	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-1155 <KUP>	
A;Cross-references: GB:AE008888; PIDN:AAL41817.1; PID:911739174; GSMPDB:GN0001	
A;Experimental source: strain C58 (Dupont)	
C;Genetics:	
A;Gene: smc	
A;Map position: circular chromosome	
Query Match 8.1%; Score 140; DB 2; Length 1155;	
Best Local Similarity 25.0%; Pred. No. 1.6;	
Matches 58; Conservatve 35; Mismatches 95; Indels 44; Gaps	
Oy 130 KQQ-----IQQFGGBNTRLHAVENLKAVNVELSEQINQL---KOLHTRLSDFG---- 175	





Copyright (c) 1993 - 2000 Compugen Ltd.

OM Protein - protein search, using sw model

Run on: August 13, 2002, 09:21:19 ; Search time 53.52 Seconds  
(without alignments)  
736.756 Million cell updates/sec

Title: US-09-673-763-8

Perfect score: 1733

Sequence: 1 MTVSTDNTSPVTASSPTF.....DDSAQPQDENQSDAGEHKIDS 355

Scoring table: BLOSUM62

Gapov 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Post-processing: Minimum Match 0%, Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802;\*

1: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1980.DAT;\*  
2: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1981.DAT;\*  
3: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1982.DAT;\*  
4: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1983.DAT;\*  
5: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1984.DAT;\*  
6: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1985.DAT;\*  
7: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1986.DAT;\*  
8: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1987.DAT;\*  
9: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1988.DAT;\*  
10: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1989.DAT;\*  
11: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1990.DAT;\*  
12: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1991.DAT;\*  
13: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1992.DAT;\*  
14: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1993.DAT;\*  
15: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1994.DAT;\*  
16: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1995.DAT;\*  
17: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1996.DAT;\*  
18: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1997.DAT;\*  
19: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1998.DAT;\*  
20: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA1999.DAT;\*  
21: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA2000.DAT;\*  
22: /SIDS1/gcdata/hold-geneseq/geneseq-emb/AA2001.DAT;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query	Match Length	DB ID	Description
1	1733	100.0	355	20 AAY32173	Chlamydia psittaci
2	200	11.5	397	20 AAY4783	Chlamydia pneumoniae
3	162.5	9.4	273	20 AAY2176	Chlamydia trachoma
4	149	8.6	612	22 AAB5546	Human protein monomer
5	145	8.4	1690	22 ABB61144	Drosophila melanogaster
6	145	8.4	1690	22 ABB51173	Drosophila melanogaster
7	143.5	8.3	1286	21 AAB43359	Human ORF X ORF3123
8	143.5	8.3	2053	22 AAB03501	Human protein kinase
9	143	8.3	1411	17 AAB0258	Nucleolar endosome
10	140.5	8.1	434	21 AAB42352	Human ORF2116
11	140	8.1	575	22 AAB47215	Human KSP-S553. S

Amino acid sequenc Human KSP. Homo sapiens sensory t Human ORFX ORF1934 Male-enhanced ant1 Mouse male enhancer Drosophila melanog Novel human diastroputative P. abyssi H. pylori secreted protein H. pylori secreted Human mitosin. Human mitochondrial amin kinase protease Human secreted protein Arabidopsis thaliana Arabidopsis thaliana Arabidopsis thaliana Arabidopsis thaliana Arabidopsis thaliana Arabidopsis thaliana Human ORFX ORF1833 Arabidopsis thaliana Novel human diastro Drosophila melanog Staphylococcus aureus Staphylococcus aureus Novel human diastro Arabidopsis thaliana Arabidopsis thaliana Arabidopsis thaliana Novel human diastro Novel human diastro

**ALIGNMENTS**

RESULT	1	AYA32173	ID	AYA32173 standard; protein; 355 AA.
XX	XX	XX	AC	AYA32173;
XX	XX	XX	DT	01-FEB-2000 (first entry)
XX	XX	XX	DE	Chlamydia psittaci infection-specific protein IncA.
XX	XX	XX	KW	IncA; Infection; vaccine; therapy; diagnosis.
XX	XX	XX	OS	Chlamydia psittaci.
XX	XX	XX	PN	WO9953948-A1.
XX	XX	XX	PD	28-OCT-1999.
XX	XX	XX	PF	20-APR-1999; 99WO-US08744.
XX	XX	XX	PR	20-APR-1998; 98US-0082438.
XX	XX	XX	PR	21-APR-1998; 98US-0082588.
XX	XX	XX	PR	22-MAY-1998; 98US-0086450.
XX	XX	XX	PI	(IYOR ) UNIV OREGON STATE.
XX	XX	XX	DR	Rockey DD, Bannantine JP;
XX	XX	XX	DR	WPI; 1999-633904/54.
XX	XX	XX	DR	N-PSDB; AAB234587.
XX	XX	XX	PT	Novel bacterial infection specific proteins for treating and diagnosing chlamydial infections -



PD	28-OCT-1998.	
XX	XX	
PF	20-APR-1999;	99WO-US08744.
XX	XX	
PR	20-APR-1998;	98US-0082438.
PR	21-APR-1998;	98US-008598.
PR	22-MAY-1998;	98US-0086450.
XX	XX	
PA	(UYOR-) UNIV OREGON STATE.	
XX	XX	
PI	Rockey DD,	Bannantine JP;
XX	XX	
DR	WPI: 1999-633904/54.	
DR	N-PSDB; AAZ34590.	
XX	XX	
PT	Novel bacterial infection - specific proteins for treating and diagnosing chlamydial infections -	
PT	Claim 5 ; Page 45-46; 56pp; English.	
XX		
PS	This sequence represents novel infection-specific protein Inca of Chlamydia trachomatis LGV-434 (serotype L2). Inca is found in the inclusion membrane of infected cells. It is associated primarily with the vegetative reticulate body form of Chlamydia rather than with the refractile elementary body form. The invention includes a vaccine directed against the reticulate body form of Chlamydia comprising 1 or more infection-specific proteins (see AAIY2210-78), including Inca, IncB and InCC; methods of using and producing such a vaccine; methods for detection of infection-specific antibodies or antigens in a biological specimen; and a method of using therapeutic agents specifically directed against infection-specific peptides, or the genes that code for such peptides, to treat chlamydial infection.	
XX		
CC		
XX		
SQ	Sequence 273 AA:	

Query Match      Best Local Similarity      Score  
                   Matches 63: *Concentric*      9.48;      162.5;  
                   Pred No. 3.2e-05;      DB 20;      Length 273;

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a 5'-end  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB82446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

Db	245	lknemenerwhigktiek1qkemadiveaarrsttrelqnldeyeknrelaemqr1k	304	Db	1177	lkethlqidderqkkfeeleeklkqaqqeqk1lgqesqtskeklteiqqs1qelqdsvkqk	1236
Qy	304	DSTNLHAVESSLIGVMYQDGAESSTVEASODDSAAQPQDE	344	Qy	240	RTRNALKELITENKTYE-----OLKADAO-LRECVRFLEKRKOE--LPEAC	285
Db	305	ekt--leakeksrtamkmqdgr--lmeeeirldyqraqde	340	Db	1237	eeiyqnleekvressiaqntklnesnvqinktsclerkdqllsqkkekqjcea	1296
				Qy	286	STLSHSIATQESTRLLKOSTTNLH---AVESLRIGTMVQDGAEASSVE-----ASQ	335
RESULT	5	ABB61144	ID ABB61144 standard; Protein; 1690 AA.	Qy	11	:    :    :    :    :    :    :    :    :    :    :    :    :	1356
DT	26-MAR-2002	(first entry)	OS	Db	1297	aklsge1qgveangdkislvkveekqatqlaqatnkelqellvksq	1356
DE	XX	Drosophila melanogaster polypeptide SEQ ID NO 10224.	XX	Qy	336	DDSAOFQDEN 345	
PN	XX	Drosophila; developmental biology; cell signalling; insecticide;	XX	Db	1357	enegnlqes 1366	
KW	XX	pharmaceutical.	OS	RESULT	6		
AC	XX	Drosophila melanogaster.	XX	ABB61173	ID ABB61173 standard; Protein; 1690 AA.		
XX	XX	W020171042-A2.	XX	XX	XX		
XX	XX	27-SEP-2001.	PD	XX	26-MAR-2002 (first entry)		
XX	XX	PF 23-MAR-2001; 2001IWO-US09241.	XX	XX	Drosophila melanogaster polypeptide SEQ ID NO 10311.		
XX	XX	PR 23-MAR-2000; 2000US-191637P.	XX	XX	Drosophila; developmental biology; cell signalling; insecticide;		
PR	XX	PR 11-JUL-2000; 2000US-0614150.	XX	XX	KW pharmaceutical.		
XX	XX	PA (PEKE ) PE CORP NY.	XX	XX	KW Drosophila melanogaster.		
PA	XX	XX Venter JC, Adams M, Li PWD, Myers EW;	PT	XX	OS WO20171042-A2.		
PI	XX	XX WPT; 2001-656860/75.	PT	XX	XX WO20171042-A2.		
PT	XX	DR N-PSDB; ABL05247.	XX	XX	XX WO20171042-A2.		
XX	XX	XX Disclosure; SEQ ID NO 10224; 21PP + Sequence Listing; English.	XX	XX	XX WO20171042-A2.		
PS	XX	PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -	XX	XX	XX WO20171042-A2.		
XX	XX	XX Disclosure; SEQ ID NO 10224; 21PP + Sequence Listing; English.	XX	XX	XX WO20171042-A2.		
XX	XX	CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL3051), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).	CC	XX	XX WO20171042-A2.		
PS	XX	CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.	CC	XX	XX WO20171042-A2.		
XX	XX	CC Sequence 1690 AA;	CC	XX	XX WO20171042-A2.		
XX	XX	CC Query Match 8.4%; Score 145; DB 22; Length 1690;	CC	XX	XX WO20171042-A2.		
XX	XX	CC Best Local Similarity 20.6%; pred. No. 0.011; Mismatches 57; Indels 90; Gaps 10;	CC	XX	XX WO20171042-A2.		
Matches	64;	CC Conservative 57; Mismatches 99; Indels 90; Gaps 10;	CC	XX	XX WO20171042-A2.		
Qy	126	VRHMKQQQQGEENTRHLTAV-----ENLKAVNVEL-----SEQINQLKL 167	CC	XX	XX WO20171042-A2.		
Db	1057	iknlqeetvtaktenlelstgtqtikldgrleitnae.qhkekmasdakradtkl 1116	CC	XX	XX WO20171042-A2.		
Qy	168	-----HTRISDFGDRLEANTGDFTLALIDFQLSL-----BE 198	CC	XX	XX WO20171042-A2.		
Db	1117	veaqvaranisaatvyleiqaeksetnhifeadminserliekvgtikee 1176	CC	XX	XX WO20171042-A2.		
Qy	199	FKSVGTVKETMLSPFEKIAQSILKE-----TFSQEAVOMMSVT 239	CC	XX	XX WO20171042-A2.		

Db	1057	iknigeetvakterneleltstgttgttkd1qerleinnaelqhkemasedaq.kiadktl	1116
Qy	168	-----HTRLSDFGDRLEANTGDFITALIAFDQLSL-----	EE 198
Db	1117	veaqyananisatraelstvleviqeksethfelemaadmsnerliekvgtlikee	1176
Qy	199	FKSYKTVTMLSFPEKLAQSLSKE-----TSQDQAVAMMSVYTF	239
Db	1177	lkethlqiderrqkfeeleeklkqaggseqqsktskeiteiqgs1qelqdsrkqk	1236
Qy	240	RTNALKALBLITENKTIV-----OLKADAQ_LREBQVRFLEBKROE---LIEAC 285	
Db	1237	eelqnleekvrsissleaqnqtklnesnvqienktsclkeqdgqliesqkkqkq.quea	1296
Qy	286	STLHSIATLQESTTLKDKSTNLH---AVESRLIGVMVQDGAESSVVE-----ASQ	335
Db	1297	aklsqeliqeqvqeqangdklslykveelykveeklqaatsqlqdaqtnkelqlvlvsq	1356
Qy	336	DBSAQPQDEN 345	
Db	1357	eneqnlages 1366	
RESULT 7			
	AAB43359	Human ORF X ORF3123 polypeptide sequence SEQ ID NO:6246.	
ID	AAB43359 standard; Protein: 1236 AA.		
XX			
AC	AAB43359;		
XX			
DT	08-FEB-2001 (first entry)		
XX			
DE	Human ORF X ORF3123 polypeptide sequence SEQ ID NO:6246.		
XX			
KW	Human; open reading frame; ORF X; detection; cytostatic; hepatotropic; vulnerability; antipsoriatric; anti-parkinsonian; nootropic; neuroprotective; anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiotonic; immunostimulant; thrombolytic; coagulant; vasodilator; hypotensive; dermatological; immunosuppressive; antiinflammatory; antiviral; antibacterial; antifungal; antiarthematic; antidiabetic; antiaemic; gene therapy; cancer; proliferative disorder; osteoarthritis; graft vs host disease; neurodegenerative disorder; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; contraceptive.		
XX			
KW	Homo sapiens.		
OS	W0200058473-A2.		
XX			
PD	05-OCT-2000.		
XX			
PF	31-MAR-2000; 2000WO-US08621.		
XX			
PR	31-MAR-1999; 99US-0127607.		
PR	02-APR-1999; 99US-0127636.		
PR	05-APR-1999; 99US-0127728.		
PR	30-MAR-2000; 2000US-0340763.		
XX			
PA	(CURA-) CURAGEN CORP.		
XX			
PA	Shimkets RA, Leach M.		
XX			
DR	WPI: 2000-602362/57.		
XX	N-PSDB; AAC7568.		
DR			
PT	Novel nucleic acids and peptides derived from open reading frame X,		
PT	useful for treating e.g. cancers, proliferative disorders,		
PT	neurodegenerative disorders and cardiovascular disease -		
XX			
Claim 11; Page 5433-5436; 5507pp; English.			

XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
 XX PPI Flanagan P, Clary D;  
 XX WPI; 2001-343950/36.  
 XX DR N/PDSB; AAS06701.  
 XX Nucleic acids encoding human kinase polypeptides, useful for preventing  
 XX PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
 XX PT neuronal-associated diseases, and microbial infections -  
 XX PS  
 XX Claim 7: Figure 2; 433pp; English.  
 XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The  
 CC novel protein kinases have been identified as members of the tyrosine  
 CC or serine/threonine kinase (PTK and STK) families. The polynucleotides  
 CC encoding protein kinases and the polypeptides may be used in the  
 CC prevention, diagnosis and treatment of diseases associated with  
 CC inappropriate kinase expression. For example, they may be used to treat  
 CC cancers (especially cancers of haematopoietic origin), cardiovascular  
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
 CC immune related diseases (e.g. rheumatoid arthritis), neurological  
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
 CC Additionally, polynucleotides encoding protein kinases may be  
 CC used for gene therapy and as DNA probes in diagnostic assays.  
 CC The protein kinase polypeptides may be used as antigens in the production  
 CC of antibodies against the protein kinases and in assays to identify  
 CC modulators of protein kinase expression and activity.  
 XX Sequence 2053 AA;  
 XX SO

Db 1023 andeivqirse-vahllreiteremdrtsqyqneairccu

RESULT 9  
AAW02258

AC	AAB42252;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	Human ORF8 ORF2116 polypeptide sequence SEQ ID NO:4232.
XX	
KW	Human; open reading frame; ORF8; detection; cytostatic; hepatotropic; vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; anticonvulsant; osteopathic; antiarthritic; immunosuppressive; cardiotonic; immunostimulant; thrombolytic; coagulant; vasotropics; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antiviral; antibacterial; antifungal; antirheumatic; antithyroid; antianemic; gene therapy; cancer; prolierative disorder; hypertension neurodegenerative disorder; osteoarthritis; graft v host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

KW Cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX Homo sapiens.  
 XX PN WO200058473-A2.  
 XX PD 05-OCT-2000.  
 XX PF 31-MAR-2000; 2000WO-US08621.  
 XX PR 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 PA (CURA-) CURAGEN CORP.  
 XX Shinkets RA, Leach M;  
 DR WPI; 2000-602362/57.  
 DR N-PSDB; AAC76561.  
 XX PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancer, proliferative disorders, neurodegenerative  
 PT disorders and cardiovascular disease -  
 XX Claim 11; Page 3424-3425; 5507pp; English.  
 CC AAC7446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytotoxic; hepatotoxic; pulmonary;  
 CC osteopathic; antiparkinsonian; nootropic; neuroprotective;  
 CC immunostimulant; cardiot; thrombolytic; coagulant; vasotropics;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, anticoagulant, bone and cartilage damage,  
 CC coagulation, to inhibit thrombosis; and as a contraceptive.  
 XX Sequence 434 AA;

Query Match 8.1%; Score 140.5; DB 21; Length 434;  
 Best Local Similarity 26.4%; Pred. No. 0.038;  
 Matches 58; Conservative 32; Mismatches 77; Indels 53; Caps 8;  
 Query 126 VRHMKQIQFQGENTRLHTAVENLKAVNVELSEQTINQLKHTRLSD-FGDRLEANTG 183  
 Db 21 ikelrdqvsgeekkilaieienksklvievneenkqkgtvlnsevleqrkvlek 80  
 Qy 184 DFTALIA----DFQSLLEEFKSVGRKVKVNLMS-----PF----- 213  
 Db 81 nrvsmilaveeyemqvnlleklirkkaesfaqemflepqgkktkppfgrqssildqq 140  
 Qy 214 ----EKLASIUKETFSQEAVQAMMSVTVERNTNALKELITENKTV--IEQLKADQ 265  
 Db 141 aldenakltqglee---eriq-hqgkvkileeqi-----enetiheihnlkqgle 187  
 Query Match 8.1%; Score 140; DB 22; Length 575;  
 Best Local Similarity 21.7%; Pred. No. 0.0062;  
 Matches 56; Conservative 56; Mismatches 100; Indels 46; Gaps 10;  
 Query 113 VITAIRDGTR-----SOVVRHMKQQIQGEENTRLHTAVENLKVANVELSEQINQLKQ 166  
 Db 298 vitalirgtrphypresktrllqdslg-grrttsi---latispasnleesttley 352  
 Qy 167 LHTRLSDFGDRLEANTG-DFTALIADFOLSLBEPK---SVGTKEVETMSPFKEQLAKQLK 221  
 Db 353 ah-raknlnkpkvnqkltkalkaikeyteeierlkrlaaareknguyiseenfrmsgk 411  
 Qy 222 ETFSQEAVQAMMSVTVERNTNALKELITENKTV--IEQLKADQ 281  
 Db 412 ltvqeegivellekieggvveeinrvtelmdnknelqckslqnktqelettqkhqqt 471



Qy	113	VITAIRDGTP-----SOYRHMKMQIQOFGEENTRLAVNVELSEQINQLQK 166
Dy	298	vitalverphvprresktrilqdsig---grtrtsi---latispaslnieetstley 352
Qy	167	LHTRLSDFGDRLEANTG-DETAIIADFOLSLBEFK----SVGTKVETMSPFEKAOSLK 221
Dy	353	ah-raknlinkpenqkltkkalkeyteeierlkrdlaarekogyiseenvrvmsg 411
Qy	222	ETPSEQEAOAMMSSTVTEIRTNLNAALKELITENKTIVIEQLKADAQLRREEQVRFLEXRKOEL 281
Dy	412	1tvgeeqiveilekigaveeeinrvrtfalmnkneidqcksdlqkgtqelettqhqhet 471
Qy	282	EENACTLSHSIATLQES-TTLLKDSTPNLHAVESSLIGVMQDGAESSTVEEASQD--- 336
Dy	472	k-----lqvyeeyitsalesteekhdaaskll-----neveettkdvsq 512
Qy	337	-----DSAQPQDENQSDDA 349
Dy	513	lhskdrrkdvqdhnaea 530
	SULT 14	
	LB96493	
LD	AAB96493 standard; Protein; 739 AA.	
XX	AC	
XX	DN	(first entry)
XX	DE	Putative sensory transduction histidine kinase and response regulator #
XX	KW	Hyperthermophilic archaeon; hyperthermophilic protein.
XX	OS	Pyrococcus abyssi.
XX	PN	FR2792651-A1.
XX	PD	27-OCT-2000.
XX	PP	21-APR-1999; 99FR-0005034.
XX	PR	21-APR-1999; 99FR-0005034.
XX	PA	(CNRS ) CNRS CENT NAT RECH SCI. (IFRE)- IFREMER INST FR RECH EXPPL MER.
XX	PI	Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
XX	PI	Quereillou J, Weissenbach J, Saurin W, Heilig R;
XX	WPI	WPI; 2001-126236/14.
	PPT	New nucleotide sequences isolated from Pyrococcus abyssi encode proteins useful in industry -
	Claim 7:	Pages 1203-1205; 165/pP; French.
	PS	
XX	XX	
CC	CC	The present invention relates to the genomic sequence of Pyrococcus abyssi (see AAB96431 and AHH1223-1) and P. abyssi proteins. P. abyssi is a hyperthermophilic archaeon, which is isolated from deep sea hydrothermal vents. The present sequence is one such P. abyssi protein. The proteins of the present invention have various potential industrial uses, since the proteins are stable at very high temperatures, some up to 110 degrees centigrade.
CC	CC	Note: This patent is in the same patent family as WO200065062, which contains additional sequences as shown in AAB99132-AAB99143, AAH15903-AHH75920 and AAG66436.
CC	CC	Sequence 739 AA;

CC AAC7446 to AAC77606 encode the proteins given in AAB00237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytosatic; hepatotropic; vulnerability;  
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressive;  
CC immunostimulant; cardiotonic; thrombolytic; coagulant; vasotropic;  
CC hypotensive; dermatological; immunosuppressive;  
CC antiinflammatory; antibacterial; antifungal; antirheumatic;  
CC antiangioid; and antianæmic. The sequences can be used for determining  
CC the presence of or predisposition to, or preventing or treating  
CC pathological conditions associated with an ORFX-associated disorder. The  
CC nucleic acids can be used to express ORFX Proteins in gene therapy  
CC vectors. The proteins and nucleic acids may be used to treat cancers,  
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
CC erythematosus, severe combined immunodeficiency (SCID), AIDS viral,  
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
CC nocturnal haemoglobinuria, thrombosis, and as a contraceptive.  
CC

Search completed: August 13, 2002, 09:23:46  
Job time: 147 sec

Result No.	Score	Query	Match	Length	DB	ID	Description
1	143.5	8.3	1286	1	CITRO_HUMAN	014578	homo sapien
2	141	8.1	944	1	NFL1_MOUSE	P32380	saccharomyces cerevisiae
3	140	8.1	1057	1	EG5_HUMAN	P52732	homo sapien
4	139	8.0	866	1	MYSP_SCHNA	P06198	schistosoma mansoni
5	139	8.0	1805	1	HSMW2_MOUSE	P47460	mycoplasma genitalium
6	138.5	8.0	1597	1	CITRO_MOUSE	P49025	mus musculus
7	136.5	7.9	1940	1	MYH3_HUMAN	P11055	homo sapien
8	136	7.9	1325	1	G160_MOUSE	P12847	rattus norvegicus
9	136	7.8	1325	1	MYH8_HUMAN	P55937	mus musculus
10	135.5	7.8	1937	1	USO1_YEAST	P13535	homo sapien
11	134	7.7	1790	1	SCPI1_MOUSE	P25386	saccharomyces cerevisiae
12	133.5	7.7	1939	1	MYH4_HUMAN	Q9y623	homo sapien
13	133	7.7	880	1	RA50_PYRAB	Q8uzc8	pyrococcus abyssi
14	132	7.6	433	1	HTR2_HALVAT	P1258	halocynthia roretzi
15	132	7.6	539	1	MYSS_HYDAT	P29922	hydra attenuata
16	132	7.6	1957	1	YDB5_SCHPO	Q10411	schizosaccharomyces pombe
17	131.5	7.6	516	1	P54_ENTFC	P13692	enterococcus faecalis
18	131.5	7.6	993	1	SCPI1_MOUSE	Q62209	mus musculus
19	131	7.6	1969	1	MYSA_CAEEL	P28444	caenorhabditis elegans
20	131	7.6	3210	1	CENF_HUMAN	P49454	homo sapien
21	131	7.6	3911	1	AKA9_HUMAN	Q99996	h. a. kinase
22	130	7.5	886	1	RA50_SULAC	Q02224	sulfolobus solfataricus
23	129.5	7.5	2663	1	CENE_HUMAN	P35749	homo sapien
24	129	7.4	1972	1	MYHB_HUMAN	P40767	bacillus subtilis
25	128.5	7.4	473	1	YNEC_BACSU	Q28641	oryctolagus cuniculus
26	128.5	7.4	1938	1	MYTH4_RABBIT	P12882	homo sapien
27	128.5	7.4	1939	1	MYH1_HUMAN	P127449	onchocerca volvulus
28	128	7.4	2022	1	ANT1_ONCOV	P12270	homo sapien
29	127.5	7.4	2349	1	TPR_HUMAN	Q9ukx3	homo sapien
30	127	7.3	1938	1	MYHD_HUMAN	P00130	dag-1
31	127	7.3	1972	1	MYHB_MOUSE	P00169	smart
32	127	7.3	1976	1	MYHA_RAT	S00036	c1; 1
33	126.5	7.3	1084	1	MYSS_RABBIT	S00033	smart

ALIGNMENTS							
<b>RESULT 1</b>							
CITRO_HUMAN							
ID	CTRO_HUMAN	STANDARD;	PRT:	1286 AA.			
AC	Q9UPZ7;						
DT	30-MAY-2000 (Rel. 39, Created)						
DT	30-MAY-2000 (Rel. 39, Last sequence update)						
DT	01-MAR-2001 (Rel. 41, Last annotation update)						
DE	Citron Protein (Rho-interacting, serine/threonine kinase 21) (Fragment).						
GN	CIT OR STRK1 OR KIAA0949.						
OS	Homo sapiens (Human)						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo;						
OX	NCBI_TaxID=9606;						
RN	[1]						
SEQUENCE FROM N.A.							
RA	Connell M.; Goela D.; Harper M.;						
RL	Submitted (Oct-1997) to the EMBL/GenBank/DDBJ databases.						
RN	[2]						
RP	SEQUENCE OF 347-1286 FROM N.A.						
RC	TISSUE-BRAIN;						
RX	MEDLINE-99z46063; PubMed-10231032;						
RA	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hiroseawa M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.; "Prediction of the coding sequences of unidentified human genes. XIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.", DNA Res. 6:63-70 (1999).						
RT	-1- FUNCTION: PUTATIVE RHO/RAC EFFECTOR THAT BINDS TO THE GTP-BOUND FORMS OF RHO AND RAC1. IT PROBABLY BINDS P21 WITH A TIGHTER SPECIFICITY IN VIVO (BY SIMILARITY).						
RT	-1- SIMILARITY: CONTAINS 1 PH DOMAIN.						
RT	-1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG BINDING DOMAIN.						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).						
CC	EMBL; AC002563; AAB71327.1; -						
DR	MIM: 605629; -						
DR	InterPro; IPR001186; CNH.						
DR	InterPro; IPR002219; DAG_PE-bind.						
DR	InterPro; IPR001849; PH.						
DR	Pfam; PR000861; RDN_repeat.						
DR	Pfam; PR00780; CINH; 1.						
DR	Pfam; PR00130; DAG_PE-bind; 1.						
DR	Pfam; PR00169; PH; 1.						
DR	SMART; SM00059; C1; 1.						
DR	SMART; SM00036; CNH; 1.						
DR	SMART; SM00233; PH; 1.						

DR	PROSITE; PS00479; DAG-PE_BIND_DOM_1; 1.	RA	Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.; Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR	PROSITE; PS50003; PH_DOMAIN; 1.	CC	-1- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIAL ROLE IN CROSSLINKING FILAMENTS OR ANCHORING OTHER MOLECULES. IT IS ESSENTIAL FOR GROWTH.
KW	Coiled coil; Phorbol-ester binding; SH3-binding.	CC	-1- SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE NUCLEUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE NUCLEOLUS.
FT	DOMAIN 1 < 1 558 COILED COIL (POTENTIAL).	CC	-1- PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.
FT	DOMAIN 1 365 561 RHO/RAC BINDING.	CC	--
FT	DOMAIN 1 509 512 POLY-LYS.	CC	--
FT	DOMAIN 1 622 670 PHOBOL-ESTER AND DAG BINDING.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).
FT	DOMAIN 1 702 822 CNH.	CC	--
FT	DOMAIN 1 851 1148 SH3-BINDING (POTENTIAL).	CC	--
SITE	1212 1217 SH3-BINDING (POTENTIAL).	CC	--
AA;	146506 MW; 498101F9EA75E85 CRC64;	CC	--
SEQUENCE	1286 AA;	CC	--
Query Match Score 8.3%; Best Local Similarity 24.3%; Matches 58; Conservative 47; Mismatches 85; Indels 49; Gaps 8;	DR	Z11582; CAA77668.1; -.	
127 RHMKQ---IQQFQEENTRHTAVENLKVANVLESEQI-----NQLKQLHTRLR 172	DR	X73297; CAA51733.1; -.	
74 RNMKPAQEMISELRLQQKFVLETLQAGKLAQNRLQEELEQKISQDHSDKNRLLLETRLR 133	DR	U28372; AAB64791.1; -.	
Db	74 DFGDRLLEANNTGDFIAQDQLSLEEFKSVGTKVETMSPFE-KLAQS--LKETF-SQ 226	DR	PIR: S26710; S26710.
Qy	173 EVSLEHEEQKLELRQLTEQLSLOQEREQLTAQQAARALESQLRQAKTELETTAE 193	DR	PIR: S34288; S34288.
Db	134 EVSLEHEEQKLELRQLTEQLSLOQEREQLTAQQAARALESQLRQAKTELETTAE 193	DR	SGD: S002764; NUF1.
FT	Coiled coil; Nuclear protein; Phosphorylation.	DR	Colled coil; Nuclear localization signal (potential).
FT	DOMAIN 164 791	DR	Nuclear localization signal (potential).
FT	DOMAIN 54 59	DR	Nuclear localization signal (potential).
FT	DOMAIN 726 731	DR	Nuclear localization signal (potential).
FT	DOMAIN 742 747	DR	Arg/Tyr-rich.
FT	DOMAIN 731 944	DR	MW; 04FAA074BB8AOBC8 CRC64;
SEQUENCE	944 AA;	DR	111781 AA;
Query Match Score 8.1%; Best Local Similarity 23.5%; Matches 61; Conservative 56; Mismatches 101; Indels 42; Gaps 8;	DR	MDLQLKOKONESEKRLKIDNLNELETKFSENGSSAKNELKMLKNAELEEELISTPKNSQ 375	
Qy	129 MKQOIQQGEENTRLHTAVENLK-----AVNVELSEQINOKLHLTRLSDFGDR 177	DR	316 MDLQLKOKONESEKRLKIDNLNELETKFSENGSSAKNELKMLKNAELEEELISTPKNSQ 375
Db	178 LEANTGDFTALIDFQLSLEEFK-----SYGVTKVETMSPFEKLQASLK ---EVFSQ- 226	DR	316 MDLQLKOKONESEKRLKIDNLNELETKFSENGSSAKNELKMLKNAELEEELISTPKNSQ 375
Qy	376 LIAKEGKLASLMA -QCLQLESKLNRQSRQLSOLRSREEELKKTNKLQDDIRIAREEIVSKD 433	DR	316 MDLQLKOKONESEKRLKIDNLNELETKFSENGSSAKNELKMLKNAELEEELISTPKNSQ 375
Db	227 EAVQAMMSVTTELRTNLNALKE---LITERKTVLEQLAD-----1EOLKADAQQLREEQVRFLE 275	DR	316 MDLQLKOKONESEKRLKIDNLNELETKFSENGSSAKNELKMLKNAELEEELISTPKNSQ 375
Qy	434 ERIDLQKVKVOLNDLFLVKKTHSEKSTDTLKDSTTNLHAVERSRLIGVMQDGAESESSTVEASQ 493	DR	316 MDLQLKOKONESEKRLKIDNLNELETKFSENGSSAKNELKMLKNAELEEELISTPKNSQ 375
Db	276 KRQKQLEEACSTLHSIATLQESTTLLKSLDKSTTNLHAVERSRLIGVMQDGAESESSTVEASQ 335	DR	316 MDLQLKOKONESEKRLKIDNLNELETKFSENGSSAKNELKMLKNAELEEELISTPKNSQ 375
Qy	494 K ---ELKEREENFYKISSEKLEDEKTTLNEKINIAAENSQNKI ---EDNSTATHIMK 546	DR	316 MDLQLKOKONESEKRLKIDNLNELETKFSENGSSAKNELKMLKNAELEEELISTPKNSQ 375
Db	336 DDSAQPDENQSDAGEBKDS 355	DR	316 MDLQLKOKONESEKRLKIDNLNELETKFSENGSSAKNELKMLKNAELEEELISTPKNSQ 375
Db	547 ENYEROLESLRKDIBEYKES 566	DR	316 MDLQLKOKONESEKRLKIDNLNELETKFSENGSSAKNELKMLKNAELEEELISTPKNSQ 375
RESULT 2	SEQUENCE FROM N.A.	DE	3
ID	NUF1_YEAST STANDARD: PRT; 944 AA.	DE	EG5_HUMAN STANDARD:
ID	NUF1_YEAST	DE	ID EG5_HUMAN STANDARD;
AC	P3.280;	DE	AC 052732; Q15716;
DT	01-OCT-1993 (Rel. 27, Created)	DE	DT 01-OCT-1996 (Rel. 34, Created)
DT	01-OCT-1993 (Rel. 27, Last sequence update)	DE	DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)	DE	DT 01-MAR-2002 (Rel. 41, Last annotation update)
GN	NUF1 protein (Spindle body spacer protein SPC110)	DE	DE Kinesin-related motor protein Eg5 (Kinesin-like spindle protein HKSP)
OS	NUF1 or SPC110 or YDR56W or D476.3.	DE	DE (Thyroid receptor interacting protein 5) (TRIP5) (Kinesin-like protein 1)
OS	Saccharomyces cerevisiae (Baker's yeast); Ascomycota; Saccharomycetes; Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Saccharomyces.	DE	DE
NCBI_TAXID=4932;	J. Cell Biol. 116:1319-1332(1992);	DE	DE
RN	SEQUENCE FROM N.A.	DE	DE
RP	SEQUENCE FROM N.A.	DE	DE
RX	STRAIN=S288C; PubMed=1541631;	DE	DE
RX	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RX	STRAIN=S288C / AB972; PubMed=7503995;	DE	DE
RA	Kilmartin J.V., Dryo S.L., Kershaw D., Finch J.T.; "A spacer protein in the <i>Saccharomyces cerevisiae</i> spindle poly body whose transcript is cell cycle-regulated.";	DE	DE
RT	"The NUF1 gene encodes an essential coiled-coil related protein that is a potential component of the yeast nucleoskeleton.";	DE	DE
RL	J. Cell Biol. 123:1175-1184(1993);	DE	DE
RN	[3] SEQUENCE FROM N.A.	DE	DE
RP	STRAIN=S288C / AB972; PubMed=1541631;	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RX	STRAIN=S288C / AB972; PubMed=7503995;	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A., Vignati D., Rifkin L., Riles L., Taich A., Trevaskis E., Vignati D.,	DE	DE
RX	SEQUENCE FROM N.A.	DE	DE
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Finch J.T.; Kershaw D., Gattung S., Greco T., Kirsten J., Kilian J., Puello A., Fulton L., Kucharski J., Hillier L., Jeter M., Kucaba T., Hallsworth K., Hawkins J., Latreille P., Le T., Johnson D., Johnston L., Langston Y., Latreille P., Miller N., Pauley A., Peluso D., Ravid E., Meneses S., Miller N., Pauley A		

RN	[1]	SEQUENCE FROM N.A., PHOSPHORYLATION SITE THR-927, AND MUTAGENESIS.	Db	298 VITALVERTPHPYRESKLTRILQDLSG - - GRTRPSI - - IATISPASLNLEETLSTLEY 352
RP		MEDLINE=9612B120; PubMed=8548803.	Qy	167 LHTRLDFGDBLEANIG - DFTALIDAEFQLSLEEPK - - - SVGTKVETMLSPFEKLAQSIIK 221
RX			Db	353 AH - RAKNLLKPEVQKLTKALIREKTEIERLKDLAARENGVYISEENRVMSSK 411
RA		"Phosphorylation by p30cdc2 regulates spindle association of human	Qy	222 EFTSQBAVQAMMSSYTELRTINNALKELENITVTEQFLKADAQREGVREFLKQEL 281
Blangy A., Lane H.A., D'Herin P., Harper M., Kress M., Nigg E.A.;		Eg5, a kinesin-related motor essential for bipolar spindle formation	RP	
RT		"Expanding the role of HseG5 within the mitotic and post-mitotic	Qy	412 LTVQEQIVELIEKVGAVEEELNRVTELFMDNKVLDQCKSDQNLQKTOLETTQKHLOET 471
RT		phases of the cell cycle." J. Cell Sci. 111:2551-2561(1998).	Db	353 AH - RAKNLLKPEVQKLTKALIREKTEIERLKDLAARENGVYISEENRVMSSK 411
RT		SEQUENCE FROM N.A.	Qy	282 BEACCTLSHSTATLQES - - TFLIKDSTTNLHAVERSLIGVMQDGAEESTVEASQD - - 336
RX		MEDLINE=9836052; PubMed=9701554;	Db	472 K - - - - - LQVKEYITSALESSTEKLHDAAASKL - - - - - NTVEPTTDVSG 512
RA		Whitehead C.M., Rattner J.B.;	Qy	337 - - - - - DSAQFQDENDSA 349
RT		"Two classes of proteins dependent on either the presence or absence	Db	513 LHSKLDRKKAVDQHNEA 530
RT		of thyroid hormone for interaction with the thyroid hormone receptor".		
RP		Mol. Endocrinol. 9:243-254(1995).		
CC	-1-	FUNCTION: MOTOR PROTEIN REQUIRED FOR ESTABLISHING A BIPOLAR SPINDLE. BLOCKING OF EG5 PREVENTS CENTROSOME MIGRATION AND ARREST	RESULT 4	
CC		CELLS IN MITOSIS WITH MONASTROBLASTIC MICROTUBULE ARRAYS.	ID	WISP_SCNA
CC	-1-	SUBUNIT: INTERACTS WITH THE THYROID HORMONE RECEPTOR IN THE PRESENCE OF THYROID HORMONE.	STANDARD;	
CC	-1-	PTM: PHOSPHORYLATED EXCLUSIVELY ON SERINE DURING S PHASE, BUT ON BOTH SERINE AND THR-927 DURING MITOSIS, SO CONTROLLING THE ASSOCIATION OF EG5 WITH THE SPINDLE APPARATUS (PROBABLY DURING EARLY PROPHASE).	PRT;	866 AA.
CC	-1-	SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC SUBFAMILY.	AC	P06198;
CC	-1-	-----	DT	01-JAN-1998 (Rel. 06, Created)
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	DT	01-JUN-1994 (Rel. 29, Last sequence update)
CC	-1-	-----	DE	16-OCT-2001 (Rel. 40, Last annotation update)
CC	-1-	-----	DEF	Paramyosin.
CC	-1-	-----	OS	Schistosoma mansoni (Blood fluke).
CC	-1-	-----	OC	Bukaryota; Metazoa; Platyhelminthes; Trematoda; Schistosomatidae.
CC	-1-	-----	OC	Rhabdophora; Eulecithophora; Revertospormate; Mediofusata; Neodermata; Trematoda; Digenea; Strigeida; Schistosomatidae.
CC	-1-	-----	OC	Schistosomatidae; Schistosoma; Schistosomatidae.
CC	-1-	-----	OX	NCBI_TAXID=6183;
CC	-1-	-----	RN	
CC	-1-	-----	RP	SEQUENCE FROM N.A.
CC	-1-	-----	RX	MEDLINE=91270282; PubMed=2052029;
CC	-1-	-----	RA	Laclette J.P., Landa A., Arcos L., Williams K., Davis A.E.,
CC	-1-	-----	RA	"Paramyosin is the Schistosoma mansoni (Trematoda) homologue of antigen B from Taenia solium (Cestoda)." Mol. Biochem. Parasitol. 44:287-296(1991).
CC	-1-	-----	RT	[1]
CC	-1-	-----	RT	SEQUENCE OF 303-742 FROM N.A.
CC	-1-	-----	RX	MEDLINE=87018840; PubMed=144;
CC	-1-	-----	RA	Ianar D.E., Pearce E.J., James S.L., Sher A.;
CC	-1-	-----	RT	"Identification of paramyosin as schistosome antigen recognized by intradermally vaccinated mice."
CC	-1-	-----	RL	Science 234:593-596(1986).
CC	-1-	-----	CC	FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
CC	-1-	-----	CC	-1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC	-1-	-----	CC	-1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC	-1-	-----	CC	-1- SIMILARITY: HIGH, TO MYOTIN HEAVY CHAINS.
CC	-1-	-----	CC	-----
CC	-1-	-----	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).
CC	-1-	-----	CC	-----
DR		EMBL: X85137; CAA59449_1; -;	DR	M55499; AAA9915_1; -;
DR		EMBL: U37426; AAA96132_1; -;	DR	M41463; AAA9914_1; -;
DR		EMBL: L03722; AAC1739_1; -;	DR	A25993; A25993.
DR		P33176; 1BG2.	DR	HSSP_P02671; 1FGZ.
DR		MINA_143760; -;	DR	InterPro_IPRO01752; kinesin.
DR		InterPro_IPRO01752; kinesin.	PFam	PF00225; kinesin_1.
DR		PRINTS: PR001380; KINESINHEAVY.	PRINTS	PR001380; KINESINHEAVY.
DR		SMART: SM00129; KISC_1.	SMART	SMART: SM00129; KISC_1.
DR		PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.	PROSITE	PS00411; KINESIN_MOTOR_DOMAIN1; 1.
KW		MOTOR protein; Microtubules; ATP-binding; Coiled coil; Mitosis;	KW	MOTOR protein; Microtubules; ATP-binding; Coiled coil; Mitosis;
KW		Phosphorylation.	FT DOMAIN	16 363 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN		364 480 COILED COIL (POTENTIAL).	FT DOMAIN	364 480 COILED COIL (POTENTIAL).
FT DOMAIN		737 764 ATP (POTENTIAL).	FT DOMAIN	737 764 ATP (POTENTIAL).
FT DOMAIN		105 112 PHOSPHORYLATION (BY CDC2).	FT DOMAIN	105 112 PHOSPHORYLATION (BY CDC2).
FT DOMAIN		927 927 T->A: NO MITOTIC PHOSPHORYLATION. NO	FT DOMAIN	927 927 T->A: NO MITOTIC PHOSPHORYLATION. NO
FT DOMAIN		BINDING TO SPINDLE APPARATUS.	FT DOMAIN	BINDING TO SPINDLE APPARATUS.
FT DOMAIN		674 676 RNS -> EL (IN REF. 2).	FT DOMAIN	674 676 RNS -> EL (IN REF. 2).
SQ		SEQUENCE 1057 AA, 119273 MW; E322F214BER1601 CRC84;	FT DISULFID	1113 VITAIRDGTP - - - - - SQVRHMKQIQOFGEENTRLTAVENLKVAVNVELSQINQIKQ 166
Query Match		Score 140; DB 1; Length 1057;	FT DISULFID	1113 VITAIRDGTP - - - - - SQVRHMKQIQOFGEENTRLTAVENLKVAVNVELSQINQIKQ 166
Best Local Similarity		Pred. No. 0.59;	FT DISULFID	1113 VITAIRDGTP - - - - - SQVRHMKQIQOFGEENTRLTAVENLKVAVNVELSQINQIKQ 166
Matches		56; Conservative 56; Mismatches 100; Indels 46; Gaps 10;	FT DISULFID	1113 VITAIRDGTP - - - - - SQVRHMKQIQOFGEENTRLTAVENLKVAVNVELSQINQIKQ 166
Qy			FT DISULFID	1113 VITAIRDGTP - - - - - SQVRHMKQIQOFGEENTRLTAVENLKVAVNVELSQINQIKQ 166

				STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMW PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYTADESIN PROTEINS IN THE MYCOPLASMAL MEMBRANE AT THE ATTACHMENT ORGANELLE (BY SIMILARITY).
FT	CONFLICT	423	428	DOYKEL -> GSSQRI (IN REF. 2).
FT	CONFLICT	430	431	SS -> KL (IN REF. 2).
FT	CONFLICT	637	637	V -> L (IN REF. 2).
FT	CONFLICT	639	639	T -> I (IN REF. 2).
FT	CONFLICT	691	691	G -> E (IN REF. 2).
FT	CONFLICT	720	720	S -> F (IN REF. 2).
SEQUENCE	866 AA;	100387 MW;	42FA5E6678176AE0	CRC64;
Query Match	8.0%	Score 139;	DB 1;	Length 866;
Best Local Similarity	21.4%	Pred. No. 0.53;	Gaps 14;	
Matches	67;	Conservative	93;	Indels 86;
Db	383 VNTLTQSQNLESNNRLRKLSLVNDLTDKNNLLENQMNQVKLSSURDANRLLTDL	442		
Qy	126 VRHMKOOFOFGENTRLHTAVEN-----LKAVNVELESEQINOLK----QLHTRLSD-	173		
Db	174 --FCDRLEANTGDETLAID-----FOLSLPEFKSVGTKVETMLSPFEKLAQSL	220		
Qy	443 EAIRSLSLEARDNIASALDAEALHDMDQKQASQAALNLHKSEMQRLLERDEELES	502		
Qy	221 KETPSOEAVGAMMSSVTYL-----RTNNAKVE-----LITENK	254		
Db	503 RKS-TTRTIPELTWITMEMVKYSELSSRLKKRYESNIATADEIQLDTANKANANLMKENK	561		
Qy	255 TVIEQLK-----ADAQLFREEQYRFLE-----KRKQ----ELEFACSTL-----SHSIA	293		
Db	562 NLSQRVKDLTEFLDERRLEERAENLNQLTEHKQLLANEEIRSTLNEFLRKIAET	621		
Qy	294 TLOESTPLLKDSTTNLHAV---ESRL---IGVMYQDGAEASTVEEAQSD-----DSAQ	340		
Db	622 ELEEAQSRVSRLTQVNTLNDKRRLEGDDIGVMQADMDDATAINAQASEPRAIRNLNEVL	681		
Qy	341 PODENOSDAGBHK	353		
Db	682 LADELROEQGNYK	694		
RESULT	5			
ID	HMW2_MYCGE	STANDARD;	PRT;	1805 AA.
AC	PA7460;			
DT	01-FEB-1996	(Rel. 33, Created)		
DT	01-FEB-1996	(Rel. 33, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Cytoadherence high molecular weight protein 2 (Cytoadherence accessory protein 2).			
DE	HMW2 OR MG218.			
OC	Bacteria; Firmicutes; Bacillus/clostridium group; Mollicutes; Mycoplasmataceae; Mycoplasma.			
OC	Mycoplasmataceae; Mycoplasma.			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RN	STRAN=ATCC 33530 / PubMed=7563993;			
RX	MEDLINE=36026446;			
RA	Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Sutton G., Kelley J.M., Fleischmann R.D., Bult C.J., Kerlavage A.R., Smith J.A., Fuhrmann J.L., Fritchman J.L., Weidman J.F., Smalik K.V., Sandusky M., Phillips C.A., Merrick J.M., Nguyen D.T., Utterback T.R., Saudek D.M., Tomb J.-F., Bott K.F., Hu P.-C., Lucier T.S., Tomb J.-F., Dougherty B.A., Bott K.F., Hutchison C.A. III, Venter J.C.; Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.; "The minimal gene complement of Mycoplasma genitalium," Science 270:397-403(1995).			
RA	[2]			
RP	SEQUENCE OF 557-659 FROM N.A.			
RP	STRAN=ATCC 33530 / G-37;			
RX	MEDLINE=9407523;			
RA	Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;			
RA	"A survey of the Mycoplasma genitalium genome by using random sequencing," J. Bacteriol. 175:7918-7930(1993).			
RT	- - FUNCTION: COMPONENT OF THE CYTOSKELETON-LIKE STRUCTURE WHICH			
CC	STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMW PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYTADESIN PROTEINS IN THE MYCOPLASMAL MEMBRANE AT THE ATTACHMENT ORGANELLE (BY SIMILARITY).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL: U39701; AAC71437.1;			
DR	EMBL: U02165; AAD12447.1;			
DR	TIGER; MG218;			
DR	InterPro; IPR003364; Seryl_tRNA_N.			
DR	Pfam; PF02403; Seryl_tRNA_N.			
KW	Cytoadherence; Structural protein; Coiled coil (potential); Coiled coil (potential); Coiled coil (potential); Coiled coil (potential); Coiled coil (potential).			
FT	DOMAIN 28 838			
FT	DOMAIN 914 159			
FT	DOMAIN 1632 1723			
FT	DOMAIN 1777 1804			
SEQUENCE	1805 AA;	216252 MW;	11D093AF173284FD CRC64;	
Query Match	8.0%	Score 139;	DB 1;	Length 1805;
Best Local Similarity	23.1%	Pred. No. 1.3;	Indels 90;	Gaps 13;
Matches	71;	Conservative	44;	Mismatches 103;
Db	304 SPFDGTTQAHVEDKVALVAKNEPKNSAKSEKLLQERELF	213		
Qy	116 AIRDGTSQVVRHMKQOIQOFGEENTRLHTAVE---NLK--AVNVYELSEQINOLQHLT	169		
Db	304 SPFDGTTQAHVEDKVALVAKNEPKNSAKSEKLLQERELF	213		
Qy	170 RLSDFGDRLLEANLGFDTALIAQFOLSL-----EFKS-VGTIVETMFLSP---F	213		
Db	359 ---FAKKHLHQNEQFEKQKODSSLKLETEYKALQHINEPKNSAKSEKLLQERELF	414		
Qy	214 EK-----LAQ-SLKETTSQEAQVAMMSVTELRNLN---ALKELITENKVIEQLK	261		
Db	415 EKRREIDTLTQSLEYHQRESSLQKLDKQNEYKQHQDNLLDQK 474			
Qy	262 ADA-----QLEREQVRF--LEKRK 278			
Db	475 VDSEATFOLKEKVAQERKELELYVKKODQKENELLFFERQLKQHQADFNEELRAQ	534		
Qy	279 QELBEACSTLHSIATLOESTTLIKDSTTNLHAVSERLIGVMVQDGAEsstVEASQDDs	338		
Db	535 QELFPAKHAIERSFIKLEDRE--KDLNTKAQQIANE-FSQLKTDKSKSADFLMLQNEY	590		
Qy	339 AQPDENQ 346			
Db	591 ENLQEQKQ 598			
RESULT	6			
CITRO_MOUSE	STANDARD;			
ID	CITRO_MOUSE			
AC	PA9025;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Citron protein (Rho-interacting, serine/threonine kinase 21).			
GN	CITR			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi.			
OC	Mammalia; Etheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.			
NCBI_TAXID	10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Brain;			
RC	TISSUE-Brain;			
RX	MEDLINE=96128238;			
RA	"A sequence from the Mycoplasma genitalium genome by using random sequencing," J. Bacteriol. 175:7918-7930(1993).			
RA	J. Bacteriol. 175:7918-7930(1993).			
CC				

RA	Madaule P., Furuyashiki T., Reid T., Ishizaki T., Watanabe G., Mori N., Narumiya S.; A novel partner for the GTP-bound forms of rho and rac."	AC P11055; Q15492; DT 01-JUL-1989 (Rel. 11, Created); DT 01-JUL-1989 (Rel. 11, Last sequence update); DT 16-OCT-2001 (Rel. 40, Last annotation update)
RL	FBBs Lett. 377:243-248(1995).	DE Myosin heavy chain, fast skeletal muscle, embryonic (Muscle embryonic myosin heavy chain) (SMHCe).
CC	-!- FUNCTION: PUTATIVE RHO/RAC EFFECTOR THAT BINDS TO THE GTP-BOUND FORMS OF RHO AND RAC1. IT PROBABLY BINDS P21 WITH A TIGHTER SPECIFICITY IN VIVO.	DE Myosin heavy chain, fast skeletal muscle, embryonic (Muscle embryonic myosin heavy chain) (SMHCe).
CC	-!- SUBUNIT: HOMODIMER (PROBABLE).	GN MYH3.
CC	-!- TISSUE SPECIFICITY: A MAJOR SIGNAL WAS OBSERVED IN TESTIS AND BRAIN, BUT IT WAS ALSO DETECTED IN THYMUS, SPLEEN, KIDNEY, HEART AND LUNG.	OS Homo sapiens (Human)
CC	-!- SIMILARITY: CONTAINS 1 PH DOMAIN.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC	-!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG BINDING DOMAIN.	OC NCBI_TaxID=9606;
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	[1]
DR	U38904; AAC2341_1; -.	RP SEQUENCE FROM N.A. MEDLINE=89233803; PubMed=2726495;
DR	MGI:1054313; C1T.	RX Eller M.S., Steedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L., Eller M.S., Steedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L., Raychowdhury M.K., Rubinstein N.A., Kelly A.M., Sarkar S.; Human embryonic myosin heavy chain cDNA. Interspecies sequence conservation of the myosin rod, chromosomal locus and isoform "Nucleotide sequence of full length human embryonic myosin heavy chain cDNA"; Nucleic Acids Res. 17:3591-3592(1989).
DR	IPR002219; DAG_PE-bind.	RX Eller M.S., Steedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L., Eller M.S., Steedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L., Raychowdhury M.K., Rubinstein N.A., Kelly A.M., Sarkar S.; Human embryonic myosin heavy chain cDNA. Interspecies sequence conservation of the myosin rod, chromosomal locus and isoform "Nucleotide sequence of full length human embryonic myosin heavy chain cDNA"; Nucleic Acids Res. 17:3591-3592(1989).
DR	IPR001849; PH.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
DR	IPR000861; REM_repeat.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
DR	Pfam: PF00789; CNH; 1.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
DR	Pfam: PF00130; DAG_PE-bind; 1.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
DR	Pfam: PF00159; PH; 1.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
DR	SMART: SM000109; CL; 1.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
DR	SMART: SM00033; PH; 1.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
DR	PROSITE: PS00479; DAG_PE_BIND_DOM_1; 1.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
DR	PROSITE: PS00081; DAG_PE_BIND_DOM_2; 1.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
KW	KW Coiled coil; Phorbol ester binding; SH3-binding.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
FT	FT DOMAIN 1 845 COILED COIL (POTENTIAL).	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
FT	FT DOMAIN 674 870 RHO/RAC BINDING.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
FT	FT DOMAIN 818 821 POLY-LYS.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
FT	FT DOMAIN 931 979 PHORBOL-ESTER AND DAG BINDING.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
FT	FT DOMAIN 1011 1131 PH.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
FT	FT DOMAIN 1160 1457 CNH.	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
FT	FT SITE 1521 1526 SH3-BINDING (POTENTIAL).	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
SO	SEQUENCE 1597 AA; 183448 MW; 7B7286C2305676DA CRC64;	RX ELLER, M.; RUBINSTEIN, N.A.; KELLY, A.M.; SARKAR, S.; CONSERVATION OF THE MYOSIN ROD, CHROMOSOMAL LOCUS AND ISOFORM "NUCLEOTIDE SEQUENCE OF FULL LENGTH HUMAN EMBRYONIC MYOSIN HEAVY CHAIN CDNA"; NUCLEIC ACIDS RES. 17:3591-3592(1989).
Query Match 8 0%; Score 138.5; DB 1; Length 1597;		
Best Local Similarity 23 8%; Pred No. 1.2;		
Matches 57; Conservative 48; Mismatches 85; Indels 49; Gaps 8;		
Qy	127 RHMKQ---IQQFGEENTHLAVENLKAVNLVEFQI-----NQLKQLHTRLS 172	CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPEATATIVE, SHOWING CYCLES OF A 28-RESIDUE REPETITION COMPOSED OF 4 HEPTAPEPTIDES, 1-CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
Ds	383 RNHKAQEENISELRLQQKFYLEAQKLQDNQRKLEQLEKISHQDHSDKSRLLELETRLR 442	CC -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
Qy	173 DFGDRLEANTGDFTLAIDLFSLEEFKQVSKVETMLSPFE-KLAQS---IKETF-SQ 226	CC -!- DEVELOPMENTAL STAGE: ABUNDANTLY PRESENT IN FETAL SKELETAL MUSCLE AND NOT PRESENT OR BARELY DETECTABLE IN HEART AND ADULT SKELETAL MUSCLE.
Ds	443 EVSIEHEEOKLELKQRLTLQSLQERSQITALQQAALESQLRQATELETTAAE 502	CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPEATATIVE, SHOWING CYCLES OF A 28-RESIDUE REPETITION COMPOSED OF 4 HEPTAPEPTIDES, 1-CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
Qy	227 EAQGAMMSSTVTELRTNLNALKE--LITENKTVIEQLKAD----- 263	CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).
Ds	503 EEIQLALTQAHDEIQRKFDALRNNSCTVITDQEQLNQLETDNAELNNQNFEYLSKOLDEA 562	CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
Qy	264 -----AQLREQVRFELERKQKLEAACSTLHSIAITLQESTTLLKDSTTNLHAYESR 317	CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
Ds	563 ANDEIVOLRS-VDHLRREITEREMQLTSQKTMALKTCTMEEQVLDEALNDL 620	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).
RESULT	7	CC or send an email to license@isb-sib.ch).
ID	MTH3_HUMAN	CC
ID	MYH3_HUMAN	CC
STANDARD:	PRT; 1940 AA.	DR X13988; CAA32167.1; -.

DR	EMBL; X13100; CAA31492; 1;	-	
DR	EMBL; X51593; CAA5942; 1;	-	
DR	EMBL; X15696; CAA33731; 1;	-	
PIR	PIR; S0490; S04090.	-	
HSSP	HSSP; P08799; 1 MMD.	-	
MIM	MIM; 160720;	-	
DR	InterPro; IPR000048; IQ,		
DR	InterPro; IPR004009; Myosin_N		
DR	InterPro; IPR02928; Myosin_tail.		
DR	InterPro; IPR01609; myosin_head.		
PFam	PF00612; IQ; 2.		
DR	PFam; PF00063; myosin_head; 1.		
DR	PFam; PF02236; Myosin_N; 1.		
DR	PFam; PF01576; Myosin_tail; 1.		
DR	PRINTS; PRO0193; MYOSINHAFY.		
DR	PRODOM; PD000355; myosin_head; 1.		
SMART	SMART; SM00015; IQ; 1.		
SMART	SMART; SM00242; MYSC; 1.		
KW	Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Alkylation; Domains in myosin family.		
FT DOMAIN	781	MYOSIN HEAD-LIKE.	
FT DOMAIN	782	811	IQ.
FT DOMAIN	840	1933	COILED COIL (POTENTIAL).
FT DOMAIN	179	186	ATP (POTENTIAL).
FT BIND	656	678	ACTIN-BINDING.
FT DOMAIN	758	772	ACTIN-BINDING.
FT MOD_RES	130	130	METHYLATION (TRI- ) (POTENTIAL).
FT MOD_RES	696	696	ALKYLATION (SH-1).
FT MOD_RES	706	706	ALKYLATION (SH-2).
FT CONFLICT	1331	1331	A -> G (IN REF. 3);
FT CONFLICT	1391	1392	KK -> RA (IN REF. 1 AND 2).
FT CONFLICT	1608	1609	SR -> RA (IN REF. 3).
FT CONFLICT	1663	1664	RG -> QT (IN REF. 2).
FT AA	1940	AA;	224035 MW; 43CA58C6A4BA1253 CRC64; SQ SEQUENCE
Query Match	7.9%	Score 136.5;	DB 1; Length 1940;
Best Local Similarity	21.4%	Pred. No. 1, 9;	
Matches 52;	Conservative	Indels 33;	Gaps 7;
QY	1.21 TPSQVVRHKQQ----1QDFGEENTRLHTA-VENLKVAV-----VELSEQTNQLKQHTR 170		
Db	1158 TSTQELNKKREAEFLKQRDLERATLQHEAMVATLRKKHADSYVAELGEQIDNLQRYK- 226		
QY	171 LSDFGDRLEANTGDFTALLADFQLSLEEFKSVGTQKVEITMSPPEKLAQSLKETFSQ --- 226		
Db	1216 -----QKLEKEKSEPKLEDDLSSSME-----SVSKRANKIEKICRDLQDSEARGK 1263		
QY	227 -EAVQAMMSYY---ELRINLNALKELTENKTIVLEQKADAQRECFVRFLEKROE 281		
Db	1264 NEEIQNSLSELTQKSRLQAFTOCEELKRQLEE 1323		
QY	282 EEAESTLHSSTATLQESTLILKDSITNLHAVESRLIGVYQDGAEESSTYEASODDQAQP 341		
Db	1324 NRKAALAHALOSSRRHDCLLREQYEQQEGKAEQLRALSANSVAQTRKYEDTAIOR 1383		
RESULT	8		
ID MYH3_RAT	STANDARD;	PRT;	1940 AA.
AC	P13847;		
DT	01-OCT-1989 (Rel. 12, Created)		
DT	01-OCT-1989 (Rel. 12, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Myosin heavy chain, fast skeletal muscle, embryonic.		
GN	MYH3.		
	Rattus norvegicus (Rat).		

Query Match 7.9%; Score 136.5; DB 1; Length 1940;  
 Best Local Similarity 21.4%; Fred. No. 1, 9;  
 Matches 52; Conservative 57; Mismatches 101; Indels 33; Gaps 7;

Qy	121	TPSQVYRHMKQQ-----IQQGEENTRLHTA-VENLKAVN----VELSEQINOLKQLHTR	170	Qy	126	VHMKQIQQ----FGEENTRL---HTAV----ENLKAVNVELSEQIN-----	162
Db	1158	TSTQIENKKREAEFLKRRDLEATTQHRTATLRRKHADSAELAQIDNIQRVK--	1215	Db	405	LQHQWYQQQLTQAVERVRQEMAHIQVGQMTQAGLBHLKLENVSLSHQTETQHR	464
Qy	171	LSDFGDRBLEANTGDFTAIDLAFOLSLSEEFKSVGTKVETMLSPFFKLQASIKETFSQ---	226	Qy	163	-----OLKQHTRLSD---FGDRBLEANTGDFTALADFOQSLEEPKSVGTPKVE	208
Db	1216	-----QKLEKEKSEKFLETDLSSSV-----SVSKSKANLKCITDLEDOLSEARGK	1263	Db	465	SIREKERIAVOLQOSIEADMLODEAFQVIREAK-----TMVBEDLORLEPEGEERQLOK	520
Qy	227	-EAVQAMMSVT---ELRNLNALKEKLLENKTVYEQLKADAQRLREQVRLEKRCBL	281	Qy	209	MLSPEFKELAOSLKE-----TESEQAVQAMMSVTELRTN	242
Db	1264	NEETQRSLSLTQTKSRLQTEAGELSRQEKEKTSVOLSRSRKAFTQIEELKROLEE	1323	Db	521	VADAAASLEQKQLEQVKLTLFQDQOALAALQEHLDVIKOLTSTQEAQKGQSLDDLHTR	580
Qy	282	EACSTLHSHTATLQESTLKLKDSTTNLHATESLIGNVQDAESSTVEEAQDDSSQP	341	Qy	243	LNLKELLENKTVYEQLKADAQRLREQVRLEBACSTPLSHSHTATLQESTTLL	302
Db	1324	NKAKLAHALQLOSSRHCDLRLREOYEEQEGKAELOLQSKANSEAVQWRTAYETDQI	1383	Db	581	YDEL-----QARLSELQREADSREDAIHFLQNKEIVALEQSAKSDEKEELDGRARRL	633
Qy	342	QDE 344		Qy	303	KDSTTNLHAVESRLIGVMQDGAEESSTVEEAQDDSAQPODENO	346
Db	1384	TEE 1386		Db	634	EEDTEE---TSGLEQLRDLAGVSNOVHQQTATLRKMQ	673
<b>SURT 9</b>							
ID	G160	MOUSE	STANDARD;	PRT;	1325	AA.	RESULT 10
AC	P55337;						MYH8_HUMAN
DT	01-NOV-1997	(Rel. 35, Created)					ID MYH8_HUMAN STANDARD;
DT	01-NOV-1997	(Rel. 35, Last sequence update)					AC P13535; O4910;
DT	15-DEC-1998	(Rel. 37, Last annotation update)					DT 01-JAN-1990 (Rel. 13, Created)
DE	Goggin-160	(Male-enhanced antigen-2) (MEA-2).					DT 15-JUL-1998 (Rel. 36, Last sequence update)
GN	GOCA3 OR MEA2.						DT 16-OCT-2001 (Rel. 40, Last annotation update)
OS	Mus musculus (Mouse).						DE Myosin heavy chain, skeletal muscle, perinatal (MyHC-perinatal).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						GN MYH8.
OC	Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.						OS Homo sapiens (Human).
OX	NCP_1_TaxID=10090;						OC Metazoa; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN	[1]						OX NCBI_TaxID=9606;
RP	SEQUENCE FROM N.A.						RN [1]
RC	STRAIN=CD-1; TISSUE=Testis;						RP SEQUENCE FROM N.A.
RX	MEDLINE=97217683; PubMed=3063644;						RC TISSUE=Skeletal muscle;
RA	Kondo M., Sutoiu S.;						RX MEDLINE=90323531; PubMed=2373371;
RT	"Cloning and molecular characterization of cDNA encoding a mouse						RA Karach-Mizrachi, I.; Feghali R., Shows T.B. Jr., Lettward L.A.;
RT	male-enhanced antigen-2 (Mea-2); a putative family of the Golgi						RT "Generation of a full-length human perinatal myosin heavy-chain-encoding cDNA." /
RL	DNA Seg. 7:1-82(1997).						RT Gene 89:289-294 (1990).
CC	-I- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR TESTIS DEVELOPMENT. PROBABLY IDENTICAL WITH THE SEROLOGICALLY DETECTABLE MALE ANTIGEN (SDM).						RN [2]
CC	-I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE FOUND IN SPERMATIDS DURING SPERMATOGENESIS. NO EXPRESSION IN LEYDIG CELLS, SPERMATOGENIA, OR SPERMATOCYTES.						RP SEQUENCE FROM N.A.
CC	-I- SIMILARITY: HIGH TO HUMAN GOGGIN-160.						RC TISSUE=Skeletal muscle;
CC	-I- CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-19 OR MET-30 IS THE INITIATOR.						RX MEDLINE=9524556; PubMed=7601129;
CC	--						RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,
CC	--						RA Steedman H.H., Rubinstein N.A.;
CC	--						RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W., Arnold H.H. /
CC	--						RT "Identification of three developmentally controlled isoforms of human myosin heavy chains." /
CC	--						RL Eur. J. Biochem. 230:1001-1006(1995).
CC	--						RN [3]
CC	--						RP SEQUENCE OF 502-1937 FROM N.A.
CC	--						RC TISSUE=Skeletal muscle;
CC	--						RX MEDLINE=9025862; PubMed=1691980;
CC	--						RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W., Arnold H.H. /
CC	--						RT "Isolation and characterization of the human perinatal myosin heavy chain." /
CC	--						RL J. Cell Biol. 108:1791-1797(1989).
CC	--						RN [4]
CC	--						RP SEQUENCE OF 1-46 FROM N.A.
DR	D78270; BAA19612.1; -.						RX MEDLINE=8934168; PubMed=2715179;
DR	HSSP_P18852; ISCG.						RA Feghali R., Lettward L.A. /
DR	MGI_MGI_96958; Golga3.						RT Molecular genetic characterization of a developmentally regulated human perinatal myosin heavy chain. /
KW	Spermatogenesis; Developmental protein.						RL J. Cell Biol. 108:1791-1797(1989).
FT	DOMAIN 201 204 POLY-ALA						RN [5]
SEQUENCE	1325 AA; 149880 MW; 3230636962C687B0 CRC64;						RP SEQUENCE OF 1-46 FROM N.A.
Query Match	Best Local Similarity 7.8%; Score 136; DB 1; Length 1325;						RA Esset K., Tidhar A., Myszkowski M. /
Matches	67; Conservative 38; Mismatches 101; Indels 78; Gaps 9;						RT Isolation and characterization of the human perinatal MHC promoter. /
							RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
							CC -I- FUNCTION: MUSCLE CONTRACTION.

-|- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
C AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

-|- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

-|- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPEITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

-|- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.

-|- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).

-|- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
-|- SIMILARITY: CONTAINS 1 IQ DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

REF ID	ACC	NAME	DEFINITION	SEQUENCE	COMMENT
CC	EMBL; M36769; AAC17185.1;				PRT; 1790 AA.
DR	EMBL; 238133; CAA86293.1;				
DR	EMBL; X51592; CAA35941.1;				
DR	EMBL; M35250; AAA36346.1;				
DR	EMBL; AF067143; AAC21557.1;				
DR	PIR; A30220; A30220.				
DR	HSSP; P08759; 1LVK.				
MIM	160741;				
InterPro	IPR00048; IQ.				
InterPro	IPR00409; Myosin_N.				
InterPro	IPR00298; Myosin_tail.				
InterPro	IPR00160; myosin_head.				
DR	PFAM; PF00612; IQ; 1.				
DR	PFAM; PF00053; myosin_head; 1.				
DR	PFAM; PF02736; Myosin_N; 1.				
DR	PFAM; PF01576; Myosin_tail; 1.				
PRINTS	PRO00193; MYOSTNHEAVY.				
DR	PRODOM; PD000355; myosin_head; 1.				
DR	SMART; SM00015; IQ; 1.				
DR	SMART; SM00242; MYSC; 1.				
DR	PROSITE; PSS0096; IQ; 1.				
DR	Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding; ATP-binding; Methylation; Alkylation; Multigene family; Catenulin binding.				
KW	MYOSIN HEAD-LIKE.				
DOMAIN	1	780	10.		
DOMAIN	781	813	COILED COIL (POTENTIAL).		
DOMAIN	842	1937			
NP-BIND	181	188	ATP.		
DOMAIN	658	680	ACTIN-BINDING.		
DOMAIN	760	774	ACTIN-BINDING.		
MOD_RES	132	132	METHYLATION (TRI-) (POTENTIAL).		
MOD_RES	698	698	ALKYLATION (SH-1) (POTENTIAL).		
MOD_RES	708	708	ALKYLATION (SH-2) (POTENTIAL).		
CONFLICT	15	15	A -> R (IN REF. 2).		
CONFLICT	970	970	E -> Q (IN REF. 1 AND 4).		
CONFLICT	1072	1072	M -> N (IN REF. 3).		
CONFLICT	1247	1247	N -> H (IN REF. 1 AND 4).		
FT	1251	1252	MC -> DGG (IN REF. 3).		
FT	CONFICT	1261	E -> G (IN REF. 1 AND 4).		
FT	CONFICT	1297	K -> Q (IN REF. 1 AND 4).		
FT	CONFICT	1377	KY -> NT (IN REF. 3).		
FT	CONFICT	1505	EN -> AH (IN REF. 1 AND 4).		
FT	CONFICT	1847	E -> D (IN REF. 1 AND 4).		
FT	CONFICT	1914	D -> H (IN REF. 2).		
FT	CONFICT	1914	A3EE2D151792E9EB CRC64;		
SEQUENCE	1937	AA;	222762 MW;		
CC	Query Match	7.8%	Score 135.5; DB 1; Length 1937;		
CC	EMBL; X54378; CAA38253.1;	-			

DR	EMBL; L03188;	AAB00143.1;	-;	
DR	U53668;	AAB66659.1;	-;	
DR	PTR: A38A45;	A38A45.	-;	
DR	HSSP; P80220;	IDIP.	-;	
DR	SGD; S0002116;	USO1.	-;	
DR	InterPro; IPR0012017;	Spectrin.	KW	
FT	Transport; Protein transport;	Golgi stack; Cytoskeleton;	Cytoskeleton;	Coiled coil.
FT	DOMAIN	1	724	GLOBULAR HEAD.
FT	DOMAIN	725	1790	COILED COIL (POTENTIAL).
FT	DOMAIN	465	1790	CHARGED (HYPER-HYDROPHILIC).
FT	DOMAIN	991	1790	DISPENSABLE FOR THE PROTEIN FUNCTION.
FT	DOMAIN	1172	1786	ASPARGLU-RICH (ACIDIC).
FT	CONFLICT	847	847	G -> E (IN REF. 2).
FT	CONFLICT	924	924	E -> K (IN REF. 2).
FT	CONFLICT	1253	1253	V -> I (IN REF. 2).
FT	CONFLICT	1319	1319	I -> V (IN REF. 2).
FT	CONFLICT	1461	1461	S (IN REF. 2).
FT	CONFLICT	1581	1581	G -> S (IN REF. 2).
FT	CONFLICT	1600	1600	I -> V (IN REF. 2).
FT	CONFLICT	1661	1661	R -> S (IN REF. 2).
FT	CONFLICT	1772	1772	D -> DEBDDE (IN REF. 2).
SEQUENCE	1790 AA;	206424 MW;	6GC2B216E9FD4819	CRC54;
Query Match	7 / 78	Score 134;	DB 1;	Length 1790;
Best Local Similarity	21.3%	Pred. No. 2.5;		
Matches	57;	Conservative	53;	Mismatches 121; Indels 36; Gaps 8;
Qy	117	IRDGTPSQVVRHMKOOIQQEENTRLHTAVENLKVAVNE-----	LSEQ	160
Db	1372	LNEG-SSTTSEYSSKINTLEDFELRLQNE- - - - - LNAKEKDNTNSLEEVLSNDELEEK	1429	
Qy	161	INQLKQHLTRISDFDRLEANTGDEFTALADFQLSLEEFK-----	SVGTKVETMSP	213
Db	1430	QNTIKSLQDELSYSDKTKITRNDKRDLESIKEQLRAAESKARVEEGLKKL	1489	
Qy	214	EKLAQSLKTF---SOEAQAMMSVTEPLTNNAKELIPTNKVTEQLADAQRE	269	
Db	1490	EE- -ESSKKAELAEKSKEMAKKLASTIENNETELKSSMETIRKSDEKLEOSKSA	--EE	1544
Qy	270	QVRFLKREKQLEEAQSTLHSATIAQLQESTTLLKQDSTNLHAEVSRLIGVM--VQDGAE	327	
Db	1545	DIKNIQHEKSDLSRINESEKDEELSKLRPEAKSGSSELEYVKQELNNAAQEKIRINAE	1604	
Qy	328	STVEEASQDDSAQPODENQSDAGEHKD	354	
Db	1605	NTVLRSKLDIERELKDQAEIKSNOE	1631	
RESULT	12			
12_HUMAN	MYHA	STANDARD;	PRT;	1939 AA.
AC	09Y23;			
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Myosin heavy chain, skeletal muscle, fetal (Myosin heavy chain 1B).			
DE	(MyHC-IIb).			
GN	MYHA.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Monilia; Fungi; Primates; Catarrhini; Hominidae; Homo.			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RC	TISSUE-Skeletal muscle;			
RX	MEDLINE=9918869; PubMed=10388558;			
RA	Weiss A.; Schiaffino S.; Leinwand L.A.;			
RA	"Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity.";			
RT	J. Mol. Biol. 290:61-75(1999).			
CC	- - FUNCTION: MUSCLE CONTRACTION.			
CC	- - SUBUNIT: MUSCLE MYOSIN IS A HOMERIC PROTEIN THAT CONSISTS OF 2			

CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1 DOMAIN: THE RODLIKE LOCATION. Thick filaments of the myofibrils.  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1 PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1 MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MERO-MYOSIN (LMM) AND 1 HEAVY MERO-MYOSIN (HMM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1 SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1 SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).  
 CC  
 DR EMBL; AF11783; AAD29949.1; -.  
 DR HSSP; P08199; 1MMD.  
 DR MIM; 160442; -.  
 DR InterPro; IPR000048; IQ.  
 DR InterPro; IPR004009; Myosin\_N.  
 DR InterPro; IPR002928; Myosin\_tail.  
 DR InterPro; IPR001600; myosin\_head.  
 DR Pfam; PF00612; IQ; 2.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 DR SMART; SM00015; IQ; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 DR Myosin\_Muscle\_protein; Colled\_coil; Thick\_filament; Actin-binding;  
 KW Calmodulin-binding; Atp-binding; Methylation; Alkylation;  
 KW Multigene family.  
 FT DOMAIN 1 784 MYOSIN HEAD-LIKE.  
 FT DOMAIN 785 814 IQ.  
 FT DOMAIN 843 1939 COILED\_COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP (POTENTIAL).  
 SQ SEQUENCE 1939 AA; 223012 MW; 40B1A1D77A47DE CRC64.

RESULT	13	ID	HTR2_HALVA	STANDARD;	PRT;	433 AA.
RA50_PYRAB	STANDARD;	PRT;	880 AA.			
ID RA50_PYRAB						
AC						
DT 16-OCT-2001 (Rel. 40, Created)						
DT 16-OCT-2001 (Rel. 40, Last sequence update)						
DT 16-OCT-2001 (Rel. 40, Last annotation update)						
DE DNA double-strand break repair rad50 ATPase.						
GN RAD50 OR PAB0812.						
OS Pyrococcus abyssi.						
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.						
OX NCBI_TaxID=29292;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC STRAIN=ATCC 29115;						
RA MEDLINE=95224074; PubMed=7707770;						
RA Seidel R., Scharf B., Gantel M., Kleine K., Oesterheld D.,						
RA Englehardt M.;						
RT "The primary structure of sensory rhodopsin II: a member of an additional retinal protein subgroup is coexpressed with its transducer, the halobacterial transducer of rhodopsin II."						
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.						
CC FUNCTION: Involved in DNA double-strand break repair (DSBR). The Rad50/mre11 complex possesses single-strand endonuclease activity and AtP-dependent double-strand-specific exonuclease activity.						
CC Rad50 provides an AtP-dependent control of mre11 by unwinding and/or repositioning DNA ends into the mre11 active site (By similarity).						
CC -1- SUBUNIT: forms a complex with mre11 (By Similarity).						
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.						
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).						
CC DR EMBL: AJ248286; CAB30131..1..;						
CC DR InterPro: IPR03593: AAA.						
CC DR InterPro: IPR03439: ABC_transporter.						
CC DR InterPro: IPR01687: ATP_GTP_A.						
CC DR InterPro: IPR01238: RecF.						
CC DR Pfam: PF00470: RecF_1.						
DR SMART: SM00352; AAA..1..;						
DR InterPro: IPR03593: AAA.						
FT NP_FIND 30 37 ATP (BY SIMILARITY).						
FT DOMAIN 144 745 COILED COIL (POTENTIAL).						
SEQUENCE 880 AA; FDB177EC7E026479 CRC64;						
Query Match Score 133; DB 1; Length 880; 7.7%;						
Best Local Similarity 23.3%; Pred. No. 1.2; 23.3%;						
Matches 48; Mismatches 89; Indels 44; Gaps 9;						
Db 465 IKKIEELKRTEERKLFF-- NLRKLEIKLUREFSYMDIAEQIKELESKLKGFNLEEL 521						
QY 126 VRHMKQQIQGEEENTRATTAVENLKAVNVELSE--- -QINQLKOLHTRLSDFG -DRL 178						
QY 179 EANTGDFTLIADFO----LSLB---- -EPKSVGTKVEMLSPPEKLAQSLK 221						
QY 522 EQKERFEGNEEFNLKGELLGLERDLKRKALEGRKLLIEEKVKAKKEELNLHQLR 581						
QY 222 ETFSQEAVQAMMSSYTELRTPNNALKELTENKTVIEQKADAOQLEEVRFLEKROEL 281						
QY 582 E-----LQFESVEELNRQLEEF -- HDKYVEAKKSESELRELNK -LEKEKTEL 629						
QY 282 EEAESTLHSSTATQESTLKLQSTNLH HAVE---- SRLIGYVQDGMESSVVE 332						
QY 630 DOAFEMIADYNEIEEKAULKDLESKFNEEYEEKRELVKLEREVSLSLARLEE 685						
Db MYS3_HYDAT ID MYS3_HYDAT STANDARD; PRT; 539 AA.						

AC P39922;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DR 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, clone 203 (Fragment).  
 OS Hydra attenuata (Hydra) (Hydra vulgaris).  
 OC Hydridae; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;  
 OC NCBITaxid:6087;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nakano M.Y., Stidwill R.P.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC -I- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE  
 CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
 CC -I- SIMILARITY: BELONGS TO THE MYOSIN HEAVY CHAIN FAMILY. STRONGEST,  
 CC TO OTHER NON MUSCLE MYOSINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).  
 CC  
 DR EMBL: L35595; AAA29216; 1;  
 DR HSSP: P24733; 1WPC.  
 DR InterPro: IPR002938; Myosin-tail.  
 PTM; PR01576;  
 DR Myosin; Myosin tail; 1.  
 KW Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 38 GLOBULAR HEAD.  
 FT DOMAIN 39 >539 RODLIKE TAIL.  
 FT DOMAIN 39 520 COILED COIL (POTENTIAL).  
 FT NON\_TER 539 539 AA; 62319 MW; 9C5AD5664060939D CRC64;

Query Match 7.68; Score 132; DB 1; Length 539;  
 Best Local Similarity 23.2%; Pred. No. 0.76;  
 Matches 62; Conservative 51; Mismatches 94; Indels 60; Gaps 10;  
 Qy 118 RDGTPSQVRHMKQQIQFGENTRLHTAVENLKVAVNVLSEQINQLKHLTRLSDFDR 177  
 Db 299 KEGLESQ-IQQLQRNIQELAKIBEELELENENKLQ-KSEL-QRKELESRIEELQDQ 352  
 Qy 178 LEANTGDFITALIADEQLSLBEFKSVGTKVETMLSPFEKLAQSU-----KE 222  
 353 LETAGGATSSQV-----EVGKKFREACNRLKEAINTANDAISAIKAKTNA 401  
 Qy 223 TFS--QEAQYAMMSVTERNTNLKELITENKTVEQLK-----ADAQLREEQVRF 273  
 Db 402 TIAEQEENAMKKAKAKLEKERSALNHNELNETKNSLDQIKKOKTNSDKNSMLERQINE 461  
 Qy 274 LEKRQELEE-----ACSTLSHSATIQUESTTLKHOSTTNJHVESRLIGYMQ 322  
 Db 462 LNSLQAQVELHSQSESNKSVSSELLANSOLSEEHNLGTATKNIKTSQL----- 515  
 Qy 323 DGAESSTYEASODDSAQPODENQSDA 349  
 Db 516 -AESKRNNEA - ESKARLKENYSSNA 538



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2002, 09:21:49 ; Search time 23.55 Seconds  
(without alignments)  
368.199 Million cell updates/sec

Title: US-09-673-763-8  
Perfect score: 1733  
Sequence: 1 MTVSDNTSPVISRASSPPF..... DDSAQQPDDNQSDAGEHKDS 355

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Maximum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:  
 1: /cgn2\_6/potodata/2/iaa/5A\_COMBO.pep:  
 2: /cgn2\_6/potodata/2/iaa/5B\_COMBO.pep:  
 3: /cgn2\_6/potodata/2/iaa/6A\_COMBO.pep:  
 4: /cgn2\_6/potodata/2/iaa/6B\_COMBO.pep:  
 5: /cgn2\_6/potodata/2/iaa/PCTUS\_COMBO.pep:  
 6: /cgn2\_6/potodata/2/iaa/backfiles1.pep:  
 \* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB	ID	Description
1	140	8.1	1057	4	US-09-541-782-10	Sequence 10, Appli
2	131	7.6	2482	1	US-08-528-284-6	Sequence 6, Appli
3	131	7.6	3248	1	US-08-353-700-1	Sequence 1, Appli
4	131	7.6	3248	1	PCT-US5-16216-1	Sequence 1, Appli
5	126	7.3	2101	1	US-08-466-390-4	Sequence 4, Appli
6	126	7.3	2101	1	US-08-470-950-4	Sequence 4, Appli
7	126	7.3	2101	1	US-08-467-761-4	Sequence 4, Appli
8	126	7.3	2101	1	US-08-495-877-4	Sequence 4, Appli
9	126	7.3	2101	2	US-08-483-924-4	Sequence 4, Appli
10	126	7.3	2101	4	US-09-452-294-1	Sequence 1, Appli
11	126	7.3	2101	5	PCT-US93-06160-4	Sequence 4, Appli
12	125.5	7.2	1184	4	US-09-541-782-2	Sequence 2, Appli
13	122.5	7.1	580	2	US-08-591-079-2	Sequence 2, Appli
14	119	6.9	976	4	US-09-104-324B-4	Sequence 4, Appli
15	118	6.8	180	1	US-08-328-554-7	Sequence 7, Appli
16	117.5	6.8	1388	2	US-08-685-776-1	Sequence 1, Appli
17	117	6.8	1312	2	US-08-592-126-148	Sequence 148, Appli
18	117	6.8	1312	2	US-08-687-080-51	Sequence 51, Appli
19	116	6.7	576	2	US-08-533-306A-2	Sequence 2, Appli
20	116	6.7	576	2	US-08-742-923A-2	Sequence 2, Appli
21	113.5	6.5	1068	4	US-09-085-199B-1	Sequence 11, Appli
22	113.5	6.5	1354	3	US-08-685-871-2	Sequence 2, Appli
23	112	6.5	546	2	US-09-067-351-1	Sequence 1, Appli
24	112	6.5	546	4	US-09-360-490-1	Sequence 1, Appli
25	112	6.5	1388	2	US-08-685-576-4	Sequence 4, Appli
26	109	6.3	896	1	US-08-095-737-2	Sequence 2, Appli
27	109	6.3	896	1	US-08-480-145-2	Sequence 2, Appli

228	109	6.3	896	2	US-08-477-389-2	Sequence 2, Appli
229	108.5	6.3	1713	3	US-08-600-982-24	Sequence 24, Appli
320	108.5	6.3	1713	5	PCT-US94-10261A-24	Sequence 24, Appli
331	108	6.2	1939	4	US-09-310-187A-1	Sequence 1, Appli
332	107.5	6.2	729	1	US-08-911-937-2	Sequence 2, Appli
333	107.5	6.2	729	2	US-08-812-533-2	Sequence 2, Appli
334	107.5	6.2	770	1	US-08-445-135-2	Sequence 2, Appli
335	107	6.2	2285	4	US-09-308-715-2	Sequence 2, Appli
336	107	6.2	414	5	PCT-US93-03077-3	Sequence 3, Appli
337	107	6.2	1093	5	PCT-US93-03077-1	Sequence 1, Appli
338	106.5	6.1	816	2	US-08-553-080A-6	Sequence 6, Appli
339	106.5	6.1	816	2	US-08-442-923A-6	Sequence 6, Appli
410	106.5	6.1	885	2	US-08-533-306A-4	Sequence 4, Appli
411	106.5	6.1	885	2	US-08-742-923A-4	Sequence 4, Appli
412	106	6.1	1388	4	US-09-572-191-2	Sequence 2, Appli
413	105	6.1	1886	4	US-08-338-105-3	Sequence 3, Appli
414	105	6.1	401	2	US-08-591-019-4	Sequence 4, Appli
415	105	6.1	608	2	US-08-336-770-1	Sequence 1, Appli

ALIGNMENTS

RESULT	US-09-541-782-10
;	Sequence 10 , Application US/09541782
;	Patent No. 6284480
;	GENERAL INFORMATION:
;	APPLICANT: Nislow, Corey
;	APPLICANT: Sakowicz, Roman
;	APPLICANT: Beraud, Christophe
;	TITLE OF INVENTION: Antifungal Assay
;	FILE REFERENCE: 1015
;	CURRENT APPLICATION NUMBER: US/09/541,782
;	CURRENT FILING DATE: 2000-04-03
;	NUMBER OF SEQ ID NOS: 10
;	SOFTWARE: FastSEQ for Windows Version 4.0
;	SEQ ID NO 10
;	LENGTH: 1057
;	TYPE: PRT
;	ORGANISM: H.sapiens
;	US-09-541-782-10
Query Match	Score 140; DB 4; Length 1057;
Best Local Simililarity	8.18;
Matches	21.7%; Pred. No. 0.00047;
Matches	Mismatches 100; Indels 46; Gaps 10
Qy	113 VITAIRDGT-----SOVVRHMKQQIOQFGEVENTRLHTAVENLKVNVELSEQINOLKQ 166
Db	298 VITALVERPHPVYRESKUTRILQDSIG -GRRTSI --IATISPASLNLEETLSTLEY 352
Qy	167 LHTRLSDFGRLEANTG-DFTALIADEQLSLEEK----SVGTKVEMUSPEEKLAQSLK 221
Db	353 AH-RAKNINKPENQKJTRKAIJKEYTEEIERLKRDIAAREKGNYIISEENFRYNSGK 411
Qy	222 ETFSQEAYQAMMSSVTTELRTNLNALKELENKTVIQOLKADAQLRQEVRLEKRKQEL 281
Db	412 LTVOEQVELKIGAEEENRVRTEFMKDNEQJCKSLQNQTQLEETTQHQIOL 471
Qy	282 EEAGSTLHSIATLQES-TTLIKDSTPNLHAEVQYQDGAESTYEEASOD--- 336
Db	472 K-----LQLVKEEYITSALESTEELKHDAAASKLL-----NTVEETKDVSG 512
Qy	337 -----DSAQPODENOSDA 349
Db	513 LHSKLDRKKAVDOHNAEA 530

## GENERAL INFORMATION:

APPLICANT: Zhu, Xueliang  
 APPLICANT: Lee, Wen Hwa  
 TITLE OF INVENTION: A NO. 5710022e. Nuclear Mitotic Phosphoprotein  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Campbell and Flores  
 STREET: 4370 La Jolla Village Drive, Suite 700  
 CITY: San Diego  
 STATE: California  
 COUNTRY: USA  
 ZIP: 92122  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 COMPUTING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/328,254  
 FILING DATE: 24-OCT-1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/141,239  
 FILING DATE: 22-OCT-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Campbell, Kathryn A.  
 REGISTRATION NUMBER: P-CJ 1191  
 REFERENCE/DOCKET NUMBER: P-CJ 1191  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 535-9001  
 TELEFAX: (619) 535-8949  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2482 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-328-254-6

TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DANN, DORMAN, HERRELL AND SKILLMAN  
 STREET: 1601 MARKET STREET, SUITE 720  
 CITY: PHILADELPHIA  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19103-2307  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 COMPUTING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/353,700  
 FILING DATE: 09-DEC-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: REED, JANET E.  
 REGISTRATION NUMBER: 36,252  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (215) 563-4100  
 TELEFAX: (215) 563-4044  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3248 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: HUMAN  
 US-08-353-700-1

Query Match 7.6%; Score 131; DB 1; Length 3248;  
 Best Local Similarity 21.9%; Pred. No. 0.016; Gaps 10;  
 Matches 64; Conservative 44; Mismatches 92; Indels 92; Gaps 10;

Qy 140 NTRLHTAVENLKVNVELSEQINOLQLH---TRLSDGDRBANTDFTLIADFQLS 195  
 Db 214 NSDLQKQCEELVQIKEEINLMAEQMHSQFAETSQRSLQDTSAHQNVVAETLSA 273  
 Qy 196 LE---EFKSYGTKVETMLSPFKELAQS-----LKETFSQEVQAMMSVTE 238  
 Db 274 LENKEKELOLNDKVTETAEIQELKKSNHLEDSLKELQLLSETSLSLEKE--MSSLIS 331  
 Qy 239 LRTNLNALKELITENKTVIEQLKADAQRLREQEVRFLEK----- 276  
 Db 332 L--NKREIEETQNSTLKE--INASLNQEMNLQKSSEFANYIDEREKSISLSLSDQY 386  
 Qy 277 -----RKOELEACSTLSSHATLQESTT---LIRDST----- 306  
 Db 387 KOEKLLLQRCHEGTGRAYEDSQKYKAQBNNSKLCLLNECTSCLCNRKNELEQLKEAF 446  
 Qy 307 -----TNLHAVESRLIGVMQDGAEESTVEEAQODDAQPODENOSDAG 350  
 Db 447 AKHQEFITKLAFABERNQNLML---ELETYQQALRSEMTDNQNNNSKSEAG 1126

RESULT 4  
 PCT-US95-16216-1  
 Sequence 1, Application PC/TUS9516216  
 GENERAL INFORMATION:  
 APPLICANT: Yen, Timothy J.  
 APPLICANT: Rattner, Jerome B.  
 TITLE OF INVENTION: Nucleic Acid Encoding a Transiently  
 EXPRESSED Kinetochoore Protein, and Methods of Use  
 TITLE OF INVENTION: Nucleic Acid Encoding a  
 NUMBER OF SEQUENCES: 4

RESULT 3  
 Sequence 1, Application US/08353700  
 Patent No. 5599319  
 GENERAL INFORMATION:  
 APPLICANT: Yen, TIMOTHY J.  
 APPLICANT: RATTNER, JEROME B.  
 TITLE OF INVENTION: Nucleic Acid Encoding a Transiently  
 EXPRESSED Kinetochoore Protein, and Methods of Use  
 TITLE OF INVENTION: Nucleic Acid Encoding a  
 NUMBER OF SEQUENCES: 4



COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/470,950  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:  
 NAME: PITCHER ESS, EDMUND R  
 REGISTRATION NUMBER: 27,829  
 REFERENCE/DOCKET NUMBER: MTP-013

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 248-7100  
 TELEFAX: (617) 248-7000

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 2101 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

US-08-470-950-4

Query Match 7.3%; Score 126; DB 1; Length 22.5%; Pred. No. 0.024;  
 Best Local Similarity 22.5%; Mismatches 97;  
 Matches 55; Conservative 48;

Matches 456 HFEERKQQEQLTDLQSSISNUS---QAKKELEQASQAH---

128 HMKQDQIQGEENTURLHTAVENLKVAVNVELSEQINQLKOLHTRLRL

184 DFTALIADQFLSIEEFKSYGTKVETMLSPFERKLAOSLKTFSQE

506 ELTNTNATIQQDDELAGIKQQAKQKA---QLAQTLOQ--QEQQ

244 NALKELITENKTYEQLKADAQLRREEQYRFLEKRKQEEACST

558 SSLKQREQQKEVAEKFQEATRDHQQL-----ATAEERDAS

304 DSTTNLHAVESRL -LGVMVQDGAEsstVE-----

612 EKAQKLELQQQLQVANEARDSAQTSVQAQREKAEILSRKVBEEL

348 DAGE 351

292 672 QVAE 675

RESULT 7  
 US-08-467-781-4 Application US/08467781  
 ; Patent No. 570596

GENERAL INFORMATION:  
 APPLICANT: TOURALY, GARY  
 APPLICANT: LIDGARD, GRAHAM P  
 TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MA  
 TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX  
 NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
 STREET: 125 HIGH STREET  
 CITY: BOSTON  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02110

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/467,781
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESO, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-467-781-4

Query Match Score 1
Best Local Similarity 27.3%; pred. N
Matches 55; Conservative 48; Mismatch

Oy 128 HMRQQIQOFGEENTRHLTAVENLKAVNVELS
Db 456 HFEERKQQLSSULIDLQSSISNLSS---QAKW
Qy 184 DFTALIAADFQLSLEEFKSVGTKVETMLSPPFV
Db 506 ELTTLNATIQQODDELAGLKQAAKEKQA---C
Qy 244 NALKELITENKTIVIEQLKADAQQLREFQVRPFT
Db 558 SSLRKQECKQLEKVAEKOETRHDQAOQ----C
Qy 304 DSTTNLHAYE8RL -IGYMVOGAEESSTVE-
Db 612 EKAAKLELQQOLQVANEARDSAQTSVTQAQ
Qy 348 DAGE 351
Db 672 QVAE 675

RESULT 8
US-08-195 487-4
; Sequence 4, Application US/08195487
; Patent No. 5783403
; GENERAL INFORMATION:
; APPLICANT: TOUKATY, GARY
; APPLICANT: LIGGARD, GRAHAM P
; TITLE OF INVENTION: NOVEL MALIGNANT
; TITLE OF INVENTION: INTERIOR NUCLEI
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA HURNITZ & THIBET
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195,48
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/901,70
; FILING DATE:

```

Page 5

ATTORNEY/AGENT INFORMATION:  
 NAME: PITCHER ESQ., EDMUND R  
 REGISTRATION NUMBER: 27, 829  
 REFERENCE/DOCKET NUMBER: MTP-013  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617/248-7000  
 TELEFAX: 617/248-7100  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2101 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-483-924-4

---

TELEFAX: (617) 248-7100  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2101 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-483-924-4

Query Match	7.3%	Score 126;	DB 2;	Length 2101;
Best Local Similarity	22.5%	Pred. No. 0	0.024;	
Matches	55:	Conservative	48.	Mismatches 0;

query Match	7.38;	Score	126,	DB	1;	Length	2101;
Best Local Similarity	22.58;	Pred. No.	0.024;				
Matches	55;	Conservative	48;	Mismatches	97;	Indels	44;
Gaps							
128	HMKQQIQFGENTRHLTAVENLAVNVELSEQINOLKOLHTRLSDFGDRLPAN	--TG	183				
456	HFEERKQQLSSLITLQSSISNLNS	--QAKEEEQAOAH	--	GARLAQVASLTS	505		
184	DFTALIADQFLSLIEFKSVGTKVMTLSPFKEQTLKETSEQEAYQAMMSVTLETRNL	243					
506	ELTLTUNATIQQDQELLAGLKQAAEKQA	--QIAQTQLO	-QEOASGLRHQEQUS	--	557		
244	NALKELITENKTVLEQLKDAQIQLREEQYRFLEXKRQKQLEEAESTLHSSTATLQESTTLLK	303					
558	SSLRKQEQQLKEVAKEQETRQDHAAQL	--	ATAAEEREPASLERDALKQLEALEK	611			
304	DSTNTHAVERSL	-IGVMVQDGASESSTIVE-----	--	EAQDDSAQPQDENQS	347		
612	EKAAKLEILQQIQLQVANEARDSAQTSVQAQREKAEELSRKVVEELQACVETARQEHEAQ	A	671				
348	DAGE	351					

JULY 9  
08-483-924-4  
Sequence 4, Application US/08483924  
Patent No. 5882876  
GENERAL INFORMATION:  
APPLICANT: TOUKATY, GARY  
APPLICANT: LILDGARD, GRAHAM P  
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE  
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX  
NUMBER OF SEQUENCES: 6

ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: 125 HIGH STREET  
CITY: BOSTON  
STATE: MA  
ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,924  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: PITCHER ESQ, EDMUND R  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: MTP-013  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000

```

; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-4833-924-4

Query Match
Best Local Similarity 7.3%; Score 126; DB 2; Length 2101;
Matches 55; Conservative 48; Mismatches 97; Indels 44; Gaps 9;

Qy 128 HMKQIQTQGFGEENTRLHTAVENLKVNVLESEQINQLKQLHTRLSDFGDRLEAN---TG 183
Db 456 HPEEEKQQLSSLTIDLOSSISNS---QAKEELEQASQAH----GARLTAQVASLTS 505

Qy 184 DFTALIADEFQSLLEEFKSVGTKVETMISPFKEPLAQSLKETSEQAVQAMMSSVTELRNL 243
Db 506 ELTTLNATIQODQDQELAGLKQAKEKQA -- QLAQTQQ - QEQASQGLRQVEQLS-- 557

Qy 244 NALKELITENKTIVIEQLKADQLRQEYRFLKRKOBLEACSTLHSIAVLOESTLLK 303
Db 558 SSLKQKEQQLKEVAEERQEAATRQDHQQL-----ATAAEREASLRDAALKQBLAEK 611

Qy 304 DSTTNQHAYEVSRL-IGMVQDGAESSTVE-----PASQDQSAQPQDENOQ 347
Db 612 EKAAKLETLQOOLQVANEARSQAQTSVTQAOREKAEELSRYEVAPEROHEAQAO 671

Qy 348 DAGE 351
Db 672 QVAE 675

RESULT 10
US-09-452-294-1
; Sequence 1, Application US/09452294
; Patent No. 6287790
; GENERAL INFORMATION:
; APPLICANT: Bissell, Mina
; TITLE OF INVENTION: UTILIZATION OF NUCLEAR STRUCTURAL PROTEINS FOR TARGETED
; TITLE OF INVENTION: THERAPY AND DETECTION OF PROLIFERATIVE AND
; TITLE OF INVENTION: DIFFERENTIATION DISORDERS
; FILE REFERENCE: IB-1454- Sequence Submittal
; Patent No. 6287790
; CURRENT APPLICATION NUMBER: US/09/452,294
; CURRENT FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: 60/110,420
; PRIOR FILING DATE: 1998-11-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-452-294-1

Query Match
Best Local Similarity 7.3%; Score 126; DB 4; Length 2101;
Matches 55; Conservative 48; Mismatches 97; Indels 44; Gaps 9;

Qy 128 HMKQIQTQGFGEENTRLHTAVENLKVNVLESEQINQLKQLHTRLSDFGDRLEAN---TG 183
Db 456 HPEEEKQQLSSLTIDLOSSISNS---QAKEELEQASQAH----GARLTAQVASLTS 505

Qy 184 DFTALIADEFQSLLEEFKSVGTKVETMISPFKEPLAQSLKETSEQAVQAMMSSVTELRNL 243
Db 506 ELTTLNATIQODQDQELAGLKQAKEKQA -- QLAQTQQ - QEQASQGLRQVEQLS-- 557

Qy 244 NALKELITENKTIVIEQLKADQLRQEYRFLKRKOBLEACSTLHSIAVLOESTLLK 303

```

RESULT 12  
 Db 558 SSIKQKEQQLKEVAKQEAQTRDQAQL---ATAAEEAREASLRERDAALKQLEALEK 611  
 Qy 304 DSTNLHAYESRL-IGVMQDGAESSTVE-----EASQDDSAQPODENOS 347  
 Db 612 EKAAKLELQQQLQVANEARSDAQTSVTAQREKAELSRSKVEELQACVETARQEQAQAA 671  
 Qy 348 DAGE 351  
 Db 672 QVAE 675  
 Db 672 QVAE 675

RESULT 11  
 PCT-US93-06160-4  
 ; Sequence 4, Application PC/TUS9306160  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NOVEL MALIGNANT CELL TYPE MARKERS OF THE  
 ; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: TESTA HURWITZ & THIBEAULT  
 ; STREET: 53 STATE STREET  
 ; CITY: BOSTON  
 ; STATE: MA USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US93/06160  
 ; FILING DATE: 19930621  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: PITCHER ESQ, EDMUND R  
 ; REGISTRATION NUMBER: 27,829  
 ; REFERENCE/DOCID NUMBER: MTP-013  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617/248-7100  
 ; TELEFAX: 617/248-7100  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2101 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; PCT-US93-06160-4

Query Match Score 126; DB 5; Length 2101;  
 Best Local Similarity 22.5%; Pred. No. 0.024; Gaps 9;  
 Matches 55; Conservative 48; Mismatches 97; Indels 44; Gaps 9;

Qy 128 HMHQIQQQGEENTRHTAVENLKAVNVELSEQINOLKQLTRLSDFGDRLEAN---TG 183  
 Db 456 HFEFEKQQLSLTDLQSSISNL---QAKEBLEQNSQAH----GARUTAQASLTS 505  
 Qy 184 DFTAIJADQSLSEEFKSFGTVETMSPFEKLAQSLKETFSQEAQAMMSVTTELRTNL 243  
 Db 506 ELTTLNATLQQDOOELAGIKQQAKEKQA---QLAQTLOQ---QEQASGLHQVEQLS--- 557

Query Match Score 125.5%; DB 4; Length 1184;  
 Best Local Similarity 19.8%; Pred. No. 0.011; Gaps 7;  
 Matches 49; Conservative 59; Mismatches 111; Indels 29; Gaps 7;

Qy 119 DGTPSQYVRHMKGQIOFGGEENTRHTAVENI----KAVNVELSEQINOLKQLTR 170  
 Db 697 NGISAAAARISEEVIGGETQLHSQNLHSFNNLQDKLKSIFTMATHLSEQNEINRLRAE 756  
 Qy 171 LSDFGDRLEAFANGDFTAIJADQSLSEEFKSFGTVETMSPFEKLAQSLKETFSQDAVQ 230  
 Db 757 LOSSRNQNLTETTKASHALA--QAEIEEHVAAEAERRTILMSQIKALVEESR---QKQFA 810  
 Qy 231 AMMSVTELTNLNALKELTENKTVEEQLKADAQRE--EQVRFLEKRKOELEEACST 287  
 Db 811 RLRAKIDGVYTERASSGNDLQEATT----QHQRQIDDWFKSEQFKDVNAASKDERTK 865  
 Qy 288 LSHSTATQESTTLLKDSTTNLHAVESRLIGYMVQDGAESSTVEAESODDSAQPQDENQS 347  
 Db 866 LQNDWEAFORNSTIRKATESVHKETVRIVDQYDDMGRQ---EALDDFVKARSON-- 920  
 Qy 348 DAGEHKDS 355  
 Db 921 --GRYRDA 926

RESULT 13  
 US-08-591-079-2  
 ; Sequence 2, Application US/08591-079  
 ; Paten No. 5972899  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Zychlinsky, Arturo  
 ; APPLICANT: Chen, Yajing  
 ; TITLE OF INVENTION: Apoptosis Induced by Shigella Ipab  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MORRISON & FOERSTER  
 ; STREET: 2000 Pennsylvania Avenue, NW  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20006-1812  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/591,079

Qy 348 DAGE 351

FILING DATE: 4:35  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Livnat, Shmuel 1  
 REGISTRATION NUMBER: 33, 949  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 887-1500  
 TELEFAX: (202) 887-0764  
 TELEX: 90-4030 MRSNOBERSH  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 580 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

Query Match 7.18; Score 122.5; DB 2; Length 580;  
 Best Local Similarity 22.7%; Pred. No. 0.0074; Indels 35; Gaps 9;  
 Matches 54; Conservative 51; Mismatches 98;

Qy 123 SOYVRHMKQTOQFGEENTRHTAVENLKVNLSEQINOLKQLHTRSDFGDRLEAN 182  
 Db 77 SQTLLIGNLQTGLERSL--TALTN-KITAWKSSQQARQKNL---EFSDRINTLL 127

Qy 183 GFTFLAIDFQSLLEKTFSYCTVEMSPFEKLAOSL-----KETFSQEAQVAMM 233  
 Db 128 SETEGLTDRYEKQINKLNKAISKDLENKINQIQTDSLNSDPESEPKKLRSREIQ--- 184

Qy 234 SSVTELTRNLNALKELTENKTVEQ--LKAQDQLREQEVRFLEKRKQELLEASTLS-H 290  
 Db 185 -----LTIKKDANKDTLTLIGORTLSHKSKNQ-LKEITDESAFANTSAS-E 234

Qy 291 STATLOESTTLLKDSSTNLNHAVESRLIGNVMQDAESSTVEEAODDSQAPQDENQD 348  
 Db 235 QLSTQOKSLLGIA-SVTQLMATTIQLGVKNNEESLKNLDAFLFQSLQESRKTEMERKS 291

RESULT 14  
 Sequence 4, Application US/09104324B  
 Patent No. 6232460  
 GENERAL INFORMATION:  
 APPLICANT: T reci, Ozlem; Sahin, Ugur; Pfreundschuh, Michael  
 TITLE OF INVENTION: Methods For Diagnosis And Treating Cancers,  
 TITLE OF INVENTION: And Methods For Identifying Pathogenic Markers In A Sample Of  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fulbright & Jaworski LLP  
 STREET: 666 Fifth Avenue  
 CITY: New York City  
 STATE: New York  
 ZIP: 10103

COMPUTER READABLE FORM:  
 COMPUTER: IBM  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/104,324B  
 FILING DATE: 25-June-1998  
 CLASSIFICATION: 435  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/892,702  
 FILING DATE: 15-July-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haunson, No. 6232460man D.  
 REGISTRATION NUMBER: 30,946  
 REFERENCE/DOCKET NUMBER: LUD 5491  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 318-3000  
 TELEFAX: (212) 752-5958  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 976 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-09-104-324B-4

Query Match 6.9%; Score 119; DB 4; Length 976;  
 Best Local Similarity 23.5%; Pred. No. 0.0033; Mismatches 80; Indels 78; Gaps 10;  
 Matches 60; Conservative 37; Mismatches 80;

Qy 139 ENTRHTAVENLKVNLSEQINOLKQLHTRSDFGDRLEANLJDFQSLBEE 198  
 Db 394 EQQRJEKNDOLKJTMELOKKSSLEEM-----TQLTNKEVELEE 435

Qy 199 FKSIVGTKVEML--SPFEKLAQSKEETPSQEAV--QAMMSVYTELRTNNAL----- 246  
 Db 436 LKKVGEKETLLEYENQFENIAEELKGQ-EQELIGLQAREKEYHDLQELQATTSBQY 494

Qy 247 ----KELTE-----NKTIV-----OLKAD--AQLEEQVRFLEKRKQ 279  
 Db 495 YSKEYKDQLTELENEKLKNTELTSHCNKLSENKELTQETSMDTELEKRNQEDINNNRKQ 554

Qy 280 ELEECSTHSIALQESTTLLKDS-----TINLHAVESRLIGMVYQDG 324  
 Db 555 E----ERMLKOIENLQETETQLRNLLEYVREELKQRDEVKCLKSEENCNLRKQVE 609

Qy 325 AESSTVEEASODDSA 339  
 Db 610 NKNKYIEELQENKA 624

RESULT 15  
 Sequence 7, Application US/08328254  
 Patent No. 5710022  
 GENERAL INFORMATION:  
 APPLICANT: Zhu, Xueliang  
 TITLE OF INVENTION: A No. 5710022 Nuclear Mitotic Phosphoprotein  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Campbell and Flores  
 STREET: 4370 La Jolla Village Drive, Suite 700  
 CITY: San Diego  
 STATE: California  
 COUNTRY: USA  
 ZIP: 92122  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentnet Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/328,254  
 FILING DATE: 24-OCT-1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/141,239  
 FILING DATE: 22-OCT-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Campbell, Cathryn A.  
 REGISTRATION NUMBER: 31,815  
 REFERENCE/DOCKET NUMBER: P-CJ 1191  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 535-9001  
 TELEFAX: (619) 535-8949  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 180 amino acids

;	TYPE: amino acid	;	Match 6.88%; Score 118; DB 1; Length 180;	;	;
;	TOPOLOGY: linear	;	Best Local Similarity 24.0%; Pred. No. 0.0033;	Gaps	7;
;	;	;	Matches 41; Conservative 41; Mismatches 47; Indels 42;	Gaps	
US-08-328-254-7					
QY	128	HMKQIQQGEENTRLHAVENLKVANVLEISQINQLQHTRLSDFGDRLEANIGDFRA	187		
Db	26	HIAEKLKERERENDSLIKVKVENLER-EIQMSEBENOEL-----	-----	61	
QY	188	LIADEFOLSLLEEFKSVGTIVKETMISPFELLAQSKTKEFQSQEAV-----QAMMSSVTTELRT	241		
Db	62	IVLDAAENSKAE-----VETLKTKQIEEMARSUK-YFDLTVLRLSEKENLTQK-QEKQG	113		
QY	242	NLNALKLTENKTVEYL-KADAQLRRE-----VRFLEKRKOELLEBACSTL	288		
Db	114	QLSEDKLSSFSKSLLEEEQAEQIKEESKPAVEMLNQNLKELNEAVAAAL	164		

Search completed: August 13, 2002, 09:24:18  
Job time: 149 sec

Copyright (c) 1993 - 2000	GenCore version 4.5	Compugen Ltd.
OM protein - protein search, using sw model		
Run on:	August 13, 2002, 09:24:04 ;	Search time 48.57 Seconds (Without alignments) 1264.427 Million cell updates/sec
Title:	US-09-673-763-8	
Perfect score:	1733	
Sequence:	1 MTVSTDNTSPVISASSPTE.....	DDSAQPODENQSAGEHKDS 355
Scoring table:	BLOSUM62	
	Gapop 10.0 , Gapext 0.5	
Searched:	2622222 seqs, 172994929 residues	
Total number of hits satisfying chosen parameters:	5622222	
Maximum DB seq length: 0		
Maximum DB seq length: 2000000000		
Post-processing: Minimum Match 0%		
Post-processing: Maximum Match 100%		
Database :	SPTREMBL_19:	Listing First 45 summaries
1: sp_archaea:*		
2: sp_bacteria:*		
3: sp_fungi:*		
4: sp_invertebrate:*		
5: sp_human:*		
6: sp_mammal:*		
7: sp_mhc:*		
8: sp_organelle:*		
9: sp_phage:*		
10: sp_plant:*		
11: sp_rhodent:*		
12: sp_virus:*		
13: sp_vertebrate:*		
14: sp_unclassified:*		
15: sp_virus:*		
16: sp_bacteriaph:*		
17: sp_archeap:*		

## ALIGNMENTS

RESULT 1  
046210 PRELIMINARY; PRT; 355 AA.  
ID Q46210;  
AC Q46210;  
DT 01-NOV-1996 (TREMBL); 01, Created  
DT 01-NOV-1996 (TREMBL); 01, Last sequence update  
DT 01-DEC-2001 (TREMBL); 19, Last annotation update  
DE INCLUSION MEMBRANE LOCALISED PROTEIN.  
GN INCIA  
OS Chamydophila caviae  
OC Bacteria; Chamydiaceae; Chamydophila.  
OX NCBI\_TAXID=83557;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GPIC;  
RX MEDLINE=9502975; PubMed=773634;  
RA Rockey D.D., Heinzen R.A., Hackstadt T.;  
RT Cloning and characterization of a Chamydia psittaci gene coding for  
a protein localized in the inclusion membrane of infected cells.;  
RL Mol. Microbiol. 15:617-626 (1995).  
DR EMBL; L35036; AAC41443.1;  
SQ SEQUENCE 355 AA; 38802 MW;

Query Match Score 100.0%; Score 1733; DB 2; Length 355;  
Best Local Similarity 100.0%; Pred. No. 4; 3e-93;  
Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MTYSTDNTSPVISASSPTEAPTSAAAGAKTATEPEGRSP 60  
Db 1 MTYSTDNTSPVISASSPTEAPTSAAAGAKTATEPEGRSP 60

Result No.	Score	Query Match	Length	DB ID	Description
1	1733	100.0	355	2 Q46210	Q46210 chamydophila
2	448	25.9	225	2 Q46210 chamydophila	052363 chamydia p
3	200	11.5	390	16 Q9282B	Q9282B chamydia p
4	185	10.7	276	16 Q9PKB8	Q9PKB8 chamydia m
5	166.5	9.6	273	2 Q9REX7	Q9REX7 chamydia t
6	164.5	9.5	273	2 Q9AMA9	Q9AMA9 chamydia t
7	164.5	9.5	273	2 Q9AM94	Q9AM94 chamydia t
8	164.5	9.5	273	16 Q9A121	Q9A121 chamydia t
9	163.5	9.4	273	2 Q99Q56	Q99Q56 chamydia t
10	162.5	9.4	273	2 Q9F7K9	Q9F7K9 chamydia t
11	162.5	9.4	273	2 Q9F7K9	Q9F7K9 chamydia t
12	162.5	9.4	273	2 Q99196	Q99196 chamydia t
13	161.5	9.3	273	2 Q9AMA7	Q9AMA7 chamydia t
14	161.5	9.3	273	2 Q9AMB2	Q9AMB2 chamydia t
15	161.5	9.3	273	2 Q9AMA6	Q9AMA6 chamydia t
16	160.5	9.3	273	2 Q9AMB1	Q9AMB1 chamydia t

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1733	100.0	355	2 Q46210	Q46210 chamydophila
2	448	25.9	225	2 Q46210 chamydophila	052363 chamydia p
3	200	11.5	390	16 Q9282B	Q9282B chamydia p
4	185	10.7	276	16 Q9PKB8	Q9PKB8 chamydia m
5	166.5	9.6	273	2 Q9REX7	Q9REX7 chamydia t
6	164.5	9.5	273	2 Q9AMA9	Q9AMA9 chamydia t
7	164.5	9.5	273	2 Q9AM94	Q9AM94 chamydia t
8	164.5	9.5	273	16 Q9A121	Q9A121 chamydia t
9	163.5	9.4	273	2 Q99Q56	Q99Q56 chamydia t
10	162.5	9.4	273	2 Q9F7K9	Q9F7K9 chamydia t
11	162.5	9.4	273	2 Q9F7K9	Q9F7K9 chamydia t
12	162.5	9.4	273	2 Q99196	Q99196 chamydia t
13	161.5	9.3	273	2 Q9AMA7	Q9AMA7 chamydia t
14	161.5	9.3	273	2 Q9AMB2	Q9AMB2 chamydia t
15	161.5	9.3	273	2 Q9AMA6	Q9AMA6 chamydia t
16	160.5	9.3	273	2 Q9AMB1	Q9AMB1 chamydia t

Qy 1 MTYSTDNTSPVISASSPTEAPTSAAAGAKTATEPEGRSP 60  
Db 1 MTYSTDNTSPVISASSPTEAPTSAAAGAKTATEPEGRSP 60

Qy 1 LLQRICLVKILIAIALEPVGIAALYCYLGIVSITPSLILMIAIMVSFVITVITARDG 120  
Db 61 LLQRICLVKILIAIALEPVGIAALYCYLGIVSITPSLILMIAIMVSFVITVITARDG 120

Qy 1 LLQRICLVKILIAIALEPVGIAALYCYLGIVSITPSLILMIAIMVSFVITVITARDG 120  
Db 61 LLQRICLVKILIAIALEPVGIAALYCYLGIVSITPSLILMIAIMVSFVITVITARDG 120

Qy 121 TPSQQVRHMKQQIQQFGEENTRLHTAVENLKAVNVELESEQINOLKOLHTLSDFGDRLEA 180  
Db 121 TPSQQVRHMKQQIQQFGEENTRLHTAVENLKAVNVELESEQINOLKOLHTLSDFGDRLEA 180

Qy	181	NTGDFTLIAIDFOLSLLEFFKSVGTKVETMISPEFEKLAOSIKETFSQEAVOAMMSEVTEL R	240	
Db	181	NTGDFTLIAIDFOLSLLEFFKSVGTKVETMISPEFEKLAOSIKETFSQEAVOAMMSEVTEL R	240	
Qy	241	TNLMALKELTENKTVIEQLKADAQRLREEQYRFLEKROELEACSTLSHISIATLQESTT	300	
Db	241	TNLMALKELTENKTVIEQLKADAQRLREEQYRFLEKROELEACSTLSHISIATLQESTT	300	
RESULT	2			
Q93263		PRELIMINARY;	PRT;	225 AA.
ID				
AC	Q53263;			
DE	01-NOV-1996 (TREMBLrel. 01, Created)			
DD	01-MAR-2001 (TREMBLrel. 01, Last sequence update)			
OS	HYPOTHETICAL 25.2 KDa PROTEIN (FRAGMENT).			
OS	Chlamydia psittaci (Chlamydophila psittaci).			
OC				
OX	NCBI_TaxID:83534;			
RN	SEQUENCE FROM N.A.			
RP	PMDB=10684935;			
RX	PMDBLINE=2015/02/25; PubMed=10684935;			
RA	Read T.D., Brunham S.C., Shen C., Gill S.R., Heidelberg J.F.,			
RA	Olinger L., Grunwood J., Davis R.W., Stephens R.S.;			
RA	"Comparative genomes of Chlamydia pneumoniae and C. trachomatis."			
RT	Nat. Genet. 21:385-389(1999).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=CWL029;			
RX	MEDLINE=992006; PubMed=10192388;			
RA	Kalman S., Matic I., Marullo P., Lammel C., Fan J., Hyman R.W.,			
RA	Olinger L., Grunwood J., Davis R.W., Stephens R.S.;			
RA	"Comparative genomes of Chlamydia pneumoniae and C. trachomatis."			
RT	Nat. Genet. 21:385-389(1999).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	PMDB=10871362;			
RX	PMDBLINE=2030349; PubMed=10871362;			
RA	Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,			
RA	Shiba T., Ishii K., Hattori M., Kubara S., Nakazawa T.;			
RA	"Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA."			
RT	from Japan and CWL029 from USA.";			
RL	Nucleic Acids Res. 28:2311-2314(2000).			
DR	A001615; AAD1839.1;			
DR	EMBL; AE001615; AAD1839.1;			
DR	EMBL; AE002217; AAF38399.1;			
DR	EMBL; APO00245; BAA98396.1;			
DR	HSSP; P0512; 1P05.			
DR	TIGR; CP0581;			
DR	Hypothetical protein; Complete proteome.			
KW	Sequence 350 AA; 45057 MW; E78B8469760A4FD4 CRC64;			
SQ	SEQUENCE			
Query Match	11.5%	Score 200;	DB 16;	Length 390;
Best Local Similarity	23.0%	Pred. No. 0_00031;		
Matches	87;	Conservative	64;	Mismatches 142; Indels 86; Gaps 13;
Qy	8 TSPVTRASSPTFGDHDGFDDNNKKLIPSTEATSSAAVGATAPEGRPLLORIC 67			
Db	2 SSPVNNTPSAPN-----IPIPAAPT-----GIPTT-----KPRSSFIERYI- 38			
Qy	68 LVKTIATIALFVGVIAALVCLYGSVIS -----TPSIL-TMLIAIMLYSFVIVITAIRGT 121			
Db	39 ---IVAKYILFAI---AATSGALTTIGLSGALTPGIGALLYTFITPSMLQLIIDSI 92			
Qy	122 PSQVYRMHQIQPFGEEN -----TRIHTAVENLKVAV----- 154			
Db	93 SGGERRLREEVSRTESENQRLTVTTTIEYKDKAQDQTLTEIAFRNENGNLKTT 152			
Qy	155 VELSHQINOLKQLTLSDFGDLRANTGDTFLAIDFQLSLEEEKS-VGTVKVTMLSP 212			
Db	153 AEDLEQSVKLSQLEFAELRINOLIQANQDAESELKLKLISGWDSKVQEINTSI- 210			
Qy	213 FEKLADSLKETTSQEAQVAMMSVYTFTRNLHAEVSLIGYVQDGAE -- -TENKTVIOLKADAQ 265			
Db	211 -----OALKVLLQEVWQEAQVHKAQEQIAQELGMHNOSTALOKSVENLVPQQ 265			
Qy	266 LREEQYRFLEKRKOELEAACSTLSHISIATLQESTTILKOSTTNHAEVESRLIGYVQDG 325			
Db	326 EESTVVEASQDSDAQDQE 344			
Qy	318 EQVTALEKMKQPAQKAEE 336			
Db	266 ALTRVGEELBESNKUSQACSLRQEIEKLAQHETSQQRIDAMLAQEONL-----A 317			
RESULT	4			
Q9Z8Z8	PRELIMINARY;	PRT;	390 AA.	
ID				
AC	Q9Z8Z8;			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	SIMILARITY TO CT119 INCA (CPJ0186 PROTEIN).			
GN	CPN0186 OR CPJ0186 OR CP0581.			
OS	Chlamydia pneumoniae (Chlamydiaceae; Chlamydophila).			
NCBI_TaxID	Chlamydiales; Chlamydiaceae; Chlamydophila.			
OX	NCBI_TaxID:83558;			
RN	SEQUENCE FROM N.A.			
RP				

DT	01-OCT-2000	(TREMBLrel. 15, Created)		
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DB		INCLUSION MEMBRANE LOCALISED PROTEIN INCA.		
GN	TC0396.			
OS	Chlamydia muridarum.			
RC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
STRAIN	=MOPN / NIGG;			
RX	MEDLINE=20150255; PubMed=10684935;			
RA	Read T.D., Brunham R.C., Shan C., Gill S.R., Heidelberg J.F.,			
RA	White R., Hickey B.K., Peterson J., Utterback T., Berry K., Bass S.,			
RA	Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,			
RA	Gwyn M., Nelson W., Debay R., Kolonay F., McClarty G., Salzberg S.L.,			
RA	Elsen J., Fraser C.M.,			
RT	"Genome sequences of Chlamydia trachomatis MoPn and Chlamydia			
RT	pneumoniae AR39.";			
RT	Nucleic Acids Res. 28:1397-1405 (2000).			
EMBL	AE002106; AAF39253.1; -.			
TIGR	TC0396; -.			
Complete proteome.				
SEQUENCE	276 AA; 30744 MW; F3C2D59A2BACD51A CRC64;			
SQ				
Query Match	10.78;	Score 185; DB 16; Length 276;		
Best Local Similarity	23.45;	Pred. No. 0.0015;		
Matches	64;	Mismatches 121;		
71;	Indels	48;		
Gaps	11;			
Qy	16 SSPTFEDHGKDFDNKKIPIPSIEAP--TSSAAVAKTATEPEGRSPLLQRICYLVKITA 73			
Db	2 TSPTLVE-----MPSCYPPTHTSSTACTKRSSSYLKPSLJETVQRVAFFVSL-- 49			
Qy	74 AIALEFVGIAALAVCLYLYGSWISTSPSLILMAIMLVSVITATIRDGTSPSQVRHMKQOL 133			
Db	50 -AIIISIGFAILGHAIIGFLIAPOVALVLIATFVITLSSIGNALYLCCKTAPLRLYELQQE 108			
Qy	134 QQGEENTRHTAVENLKVANVLESEQINQLQHLTRLSFGDRLEANFTGDFTAIDAQ 193			
Db	109 ASLKEVNFL-----LKSQKEF---LGISKDFATTSKDLSD---VSDFHNLQDFQ 154			
Qy	194 LSLEEFKSYGTRVETMLSPPEKLAQSLKTFQSQAVQAMMSSTELRNMLNALKEL--- 249			
Db	155 SSHQGF-----EDLDKYKNSADLRLQFSQTQVSKLSTLSKEEIKELVPLTEEV 207			
Qy	250 --ITENKTYTEQLKADAQ-LREEQYRELERKQELEAACSTLSHSIAT-LOESTTLKDS 305			
Db	208 RRLKENKDLKLTKIVQDLDIHKD---LRAEINNLSQLASKTLSEQASQEEENEKLANT 263			
Qy	306 TTNL 309			
Db	264 TKAL 267			
RESULT	5			
ID	Q9RFX7	PRELIMINARY;	PRT;	273 AA.
AC	Q9RFX7;			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)		
DE	INCLUSION MEMBRANE PROTEIN A.			
GN	INCA.			
OS	Chlamydia trachomatis.			
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.			
NCBI_TAXID	=813;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
STRAIN	=D(S) 2923;			
MEDLINE	=20072706; PubMed=10603409;			
RA	Suchland R.J., Rockey D.D., Bannantine J.P., Stamm W.E.;			
RA	"Isolates of Chlamydia trachomatis that occupy nonfusogenic inclusions			
RT	Infect. Immun. 69:4654-4656(2001);			
RC	Strain=M9346, MT9291, AND MT9336;			
RX	PubMed=11402010;			
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van den Brule A.J.C., More S.A., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the inca 147T mutation."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL	[3]			
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the inca 147T mutation."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the inca 147T mutation."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				
RN				
SEQUENCE FROM N.A.				
STRAIN	=IC-CA08;			
RC				
RA	Pannekoek Y., van der Ende A., Eijk P.P., van Marle J., de Witte M.A.,			
RA	Ossewaarde J.M., van Marle J., More S.A., van den Brule A.J.C., Dankert J.;			
RT	"Normal Inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates that Occupy nonfusogenic inclusions."			
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL				

DR	EMBL; AF327003; AAG61100.1; - . 61FA8F2085701F58 CRC64;	Query Match 9.5%; Score 164.5; DB 2; Length 273; Best Local Similarity 22.8%; Pred. No. 0_023; Mismatches 59; Indels 59; Gaps 10; Matches 67; Conservative	Db	214 EELKTIIRDNLREIGQLSOLSKLTSQLTQS1ALQRKESSDLSQIRETLLSPRSKAS 267
Qy	32 IPISTEAPSSAAAGAKTALEPEGRSPILLORICLYVKAIAIAFLVFGTAIALVCLYLG 91	RESULT 8	084121 PRELIMINARY; PRT; 273 AA.	
Db	8 VTPPSPPPAPSYSANR-----POPSLMRDKIKKAAIASLILITGTLALLGHLVG 58	ID 084121; AC 084121; DT 01-NOV-1998 (TREMBLrel. 08, Created) DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update) DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
Qy	92 SVI7PSLMLATMLYSFVIVITAIRDGTPSOVVR---HMKQIQQFGEENTRLHTAV 147	DE INCLUSION MEMBRANE PROTEIN A.		
Db	59 FLTPQTIVLLAFLTISLA-----GNAYLQKTANLHLYQDLQR-----EV 100	GN INCA OR CPT119. OS Chlamydia trachomatis. OC Bacteria; Chlamydiaceae; Chlamydiae. OC NCBI_TAXID=813; OX RN		
Qy	148 ENIKAVNVELESEQINOLKOLHPLSDFGDRLEANTGDFTLADEFQSLSEEFKSVGTKE 207	SEQUENCE FROM N.A.		
Db	101 GSLEKINFMLS-VLOKEFLH----LSKEFATSKSOLASQDFYSCLOQFRDNFKGFE 153	RP STRAIN=D/TW-3/CX; RX MEDLINE=99000809; PubMed=9784136; RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L., RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., RA Davis R.W.; RT Genome sequence of an obligate intracellular pathogen of humans: RA Chlamydia trachomatis."; RL Science 282:754-759(1998).		
Qy	208 TMISPFKEKLAQSKETTSQETEPAVAMMSSVTELTINNALKELITENK-----TVI 257	RN [1]		
Db	154 SLLDEYKNSTEEMRKLFQSOETIADLKGSVASYLREEIRFLPLAEVRLAHNQSLTVVI 213	SEQUENCE FROM N.A.		
Qy	258 EQLKA-DAQIREE-QVRPLEK-----TKOFLEEACTLSHSIAATLQUEST 300	RC Viratayosin W., Rockey D.D., Suchland R.J., Stamm W.E.; RT "Diverse Mutations in inca Amplified from Clinical Chlamydia trachomatis Isolates that Occupy Nonfusogenic Inclusions."; RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.		
Db	214 EELKTIIRDNLREIGQLSOLSKLTSQLTQS1ALQRKESSDLSQIRETLLSPRSKAS 267	RN [2]		
Qy	99 AM94 7 PRELIMINARY; PRT; 273 AA.	SEQUENCE FROM N.A.		
Db	AC 09AM94; DT 01-JUN-2001 (TREMBLrel. 17, Created) DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update) DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	RP STRAIN=MT9301; RC Viratayosin W., Rockey D.D., Suchland R.J., Stamm W.E.; RT "Diverse Mutations in inca Amplified from Clinical Chlamydia trachomatis Isolates that Occupy Nonfusogenic Inclusions."; RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.		
Qy	111 7 PRELIMINARY; PRT; 273 AA.	RN [3]		
Db	AC 09AM94; DT 01-JUN-2001 (TREMBLrel. 17, Created) DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update) DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	SEQUENCE FROM N.A.		
Qy	111 7 PRELIMINARY; PRT; 273 AA.	RC STRAIN=SA-1, APACHE-2, IOL-238, UW-36, 2B, 10A, AND 11A; RA Pannekoek Y., van der Ende A., Eijk P.P., van Marie J., de Witte M.A., RA Ossewaarde J.M., Morre S.A., van den Brule A.J.C., Dankert J.; RT "Normal inca expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the inca 14/77 mutation."; RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.		
Db	AC 09AM94; DT 01-JUN-2001 (TREMBLrel. 17, Created) DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update) DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DR EMBL; AE001286; AAC67710.1; DR EMBL; AF279346; AAC31466.1; DR EMBL; AF327329; AAC11233.1; DR EMBL; AF326992; AAC61089.1; DR EMBL; AF326994; AAC61091.1; DR EMBL; AF326998; AAC61095.1; DR EMBL; AF327002; AAC61099.1; DR EMBL; AF327004; AAC61101.1; DR EMBL; AF327006; AAC61103.1; DR EMBL; AF327012; AAC61109.1; DR EMBL; AF327014; AAC61111.1; DR EMBL; AF327015; AAC61112.1; KW Complete proteome: SQ SEQUENCE 273 AA; 30327 MW; 6945AE7EB8BBBF3 CRC64;		
Qy	111 7 PRELIMINARY; PRT; 273 AA.	DR EMBL; AF327329; AAC11233.1; DR EMBL; AF326992; AAC61089.1; DR EMBL; AF326994; AAC61091.1; DR EMBL; AF326998; AAC61095.1; DR EMBL; AF327002; AAC61099.1; DR EMBL; AF327004; AAC61101.1; DR EMBL; AF327006; AAC61103.1; DR EMBL; AF327012; AAC61109.1; DR EMBL; AF327014; AAC61111.1; DR EMBL; AF327015; AAC61112.1; KW Complete proteome: SQ SEQUENCE 273 AA; 30327 MW; 6945AE7EB8BBBF3 CRC64;		
Db	148 ENIKAVNVELESEQINOLKOLHPLSDFGDRLEANTGDFTLADEFQSLSEEFKSVGTKE 207	Query Match 9.5%; Score 164.5; DB 2; Length 273; Best Local Similarity 22.8%; Pred. No. 0_023; Mismatches 59; Indels 59; Gaps 10; Matches 67; Conservative	Db	Query Match 9.5%; Score 164.5; DB 16; Length 273; Best Local Similarity 22.8%; Pred. No. 0_023; Mismatches 59; Gaps 10; Matches 67; Conservative
Qy	101 GSLEKINFMLS-VLOKEFLH----LSKEFATSKSOLASQDFYSCLOQFRDNFKGFE 153	Db	Db	Db
Qy	208 TMISPFKEKLAQSKETTSQETEPAVAMMSSVTELTINNALKELITENK-----TVI 257	32 TIPISEAPTSAAAGAKTALEPEGRSPILLORICLYVKAIAIAFLVFGTAIALVCLYLG 91	Qy	Qy
Db	154 SLLDEYKNSTEEMRKLFQSOETIADLKGSVASYLREEIRFLPLAEVRLAHNQSLTVVI 213	-----TVI 257	8 VTPPSPPPAPSYSANR-----POPSLMRDKIKKAAIASLILITGTLALLGHLVG 58	Db
Qy	258 EQLKA-DAQIREE-QVRPLEK-----TKOFLEEACTLSHSIAATLQUEST 300	-----TVI 257	8 VTPPSPPPAPSYSANR-----POPSLMRDKIKKAAIASLILITGTLALLGHLVG 58	Qy



Page 6

RA	Ossewaarde J.M., van Den Brule A.J.C., Morre S.A., Dankert J.;	QY	208 TMLSPFEKLAQSIIKETESQAVQAMSSVTLRT-----
RT	Normal InCA expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the inca I47T mutation.;	Db	154 SLDEYRNSTEENMRKLFSEQLTADLKSVASLREIRELTPLAEVCRALHNRESLTAAI 213
RL	Infect. Immun. 69:4454-4556 (2001).	Db	154 SLDEYRNSTEENMRKLFSEQLTADLKSVASLREIRELTPLAEVCRALHNRESLTAAI 213
DR	EMBL; AF327005; AAG1102.1; -;	QY	251 TENKTIVELQLKAD---AQLRREEQVREPLEKIRQELBACSTLSHSTATLOESTT 300
SQ	SEQUENCE 273 AA; 30387 MW; 6945AE7E9B8AEFB3 CRC64;	Db	214 EELKTIDSLRDEIGOLSQLSKLTSQLQRKESSDLCSQIRELSSPRKSAS 267
Query Match	9.3%; Score 161.5; DB 2; Length 273;		
Best Local Similarity	22.8%; Pred. No. 0.05;		
Matches	67; Conservative 58; Mismatches 110; Indels 59; Gaps 10;		
		Search completed: August 13, 2002, 09:30:16	
		Job time: 372 sec	
QY	32 IIPISTIEPTSSAAVGAKTATEPEGRSPLLORICLYVKVIIIAAIALEFVGTAALVCLYLG 91		
Db	8 VTPPSPPAPSYSANRV-----POPSLMDKIKKIAASLILIGTGFIALIGHYG 58		
QY	92 SVISTPSSLMLIAIMLYSFTVITIAIRDGTSPQVVR----HMKQIQQFGEENTRLHTAV 147		
Db	59 FLIAPQTIVVLALFITSLA-----GNALYLOKTANLHYQDLQR-----EV 100		
QY	148 ENLKAVNVELSEQINOLKQLHTRPLSDFGDRBANTGDETALIAFDOLSLBEEFKSYGTKE 207		
Db	101 GSLEKINEIFMLS - VLOREFLH -----LSKEFATTSDKLDSAYSQDFYVSCLQFRDNWKGE 153		
QY	208 TMLSPFEKLAQSIIKETESQAVQAMSSVTLRT-----		
Db	154 SLDEYRNSTEENMRKLFSEQLTADLKSVASLREIRELTPLAEVCRALHNRESLTAAI 213		
QY	258 EQLKA -DQQLRBE -QVRFLEX-----RKQELPEACSTLSHSTATLOESTT 300		
Db	214 EELKTIDSLRDEIGOLSQLSKLTSQLQRKESSDLCSQIRELSSPRKSAS 267		
RESULT	15		
Q9AMA6	PRELIMINARY; PRT; 273 AA.		
ID	Q9AMA6		
AC	Q9AMA6;		
DT	01-JUN-2001 (Tremblrel. 17, Created)		
DT	01-DEC-2001 (Tremblrel. 19, Last sequence update)		
DE	INCLUSION MEMBRANE PROTEIN A.		
GN	INCA.		
OS	Chlamydia trachomatis.		
OC	Bacteria; Chlamydiidales; Chlamydiaceae; Chlamydia.		
OX	NCBI_TaxID=813;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=440-L;		
RX	MEDLINE=21295121; PubMed=11402010;		
	Pannenkoek Y., van Der Ende A., Eljik P.P., van Marle J., de Witte M.A.,		
	"Normal InCA expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the inca I47T mutation.";		
	Infect. Immun. 69:4454-4556 (2001).		
	EMBL; AF327005; AAG1102.1; -;		
	SEQUENCE 273 AA; 30265 MW; 2267957C555B572B CRC64;		
Query Match	9.3%; Score 161.5; DB 2; Length 273;		
Best Local Similarity	21.4%; Pred. No. 0.035;		
Matches	63; Conservative 61; Mismatches 111; Indels 59; Gaps 8;		
QY	32 IIPISTIEPTSSAAVGAKTATEPEGRSPLLORICLYVKVIIIAAIALEFVGTAALVCLYLG 91		
Db	8 VTPPSPPAPSYSANRV-----POPSLMDKIKKIAASLILIGTGFIALIGHYG 58		
QY	92 SVISTPSSLMLIAIMLYSFTVITIAIRDGTSPQVVR----HMKQIQQFGEENTRLHTAV 147		
Db	59 FLIAPQTIVVLALFITSLA-----GNALYLOKTANLHYQDLQR-----EV 100		
QY	148 ENLKAVNVELSEQINOLKQLHTRPLSDFGDRBANTGDETALIAFDOLSLBEEFKSYGTKE 207		
Db	101 GSLEKINEIFMLS - VLOREFLH -----LSKEFATTSDKLDSAVSQDFYSCLGFRDNWKGE 153		

